

Db 1165 pProLeuGlnTyrMetArgAlaAspGluAlaGly----- 1177  
 Qy 176 GCCTAGGAATCAGCCAGCGCCCATTTCTGCCAGCCCTTTGGTGGCGGTCCA----- 125  
 Db 1178 -----GlyLeuArgProHisAspG1 1184  
 Qy 124 -----GCTTCTCAGCCCATGCTCAACACCTGCTGCTGGGGCACCCTCAGTGGGG 75  
 Db 1184 uGluValGluAlaThrLeuLysSerLeuAsn-----As 1195  
 Qy 74 ACACGTCTCATCTACTCAGATCTCTGGCGAGCGCGGCTGTACCCGAGCCAGCGCGCT 15  
 Db 1195 nGlnIleGluSerIleArgSer-ProGluGlySerArgLysAsnPro-----AlaArgThrC 1214  
 Qy 14 GCAGG 10  
 Db 1214 ysarg 1215  
 RESULT 20  
 CGH06C  
 collagen alpha 1(II) chain precursor [validated] - human  
 N:Alternate names: procollagen alpha-1(II) chain  
 C:Contains: chondrocalcin; collagen alpha 1(II) chain precursor splice form 1; collagen  
 C:Species: Homo sapiens (man)  
 C:Date: 28-May-1986 #sequence revision 01-Sep-1995 #text change 08-Dec-2000  
 C:Accession: A38513; S06715; S24270; A24828; S06496; A35428; A30147; A33116; S64674; S63  
 7250; I37251; I37252; I37253; I37254; I55338; I55335; I61910  
 R:Ryan, M.C.; Sieraski, L.J.  
 Genomics 8, 41-48, 1990  
 A:Title: The human type II procollagen gene: identification of an additional protein-cod  
 A:Reference number: A38513; MUID:91184811; PMID:2081599  
 A:Accession: A38513  
 A:Molecule type: DNA  
 A:Residues: 1-103 <RYA>  
 R:Su, M.W.; Lee, B.; Ramirez, F.; Machado, P.; Horton, W.  
 Nucleic Acids Res. 17, 9473, 1989  
 A:Title: Nucleotide sequence of the full length cDNA encoding for human type II procolla  
 A:Reference number: S06715; MUID:90067946; PMID:2587267  
 A:Accession: S06715  
 A:Molecule type: mRNA  
 A:Residues: 1-28, 'R', '99-1487 <SU2>  
 A:Cross-references: EMBL:X16468; NID:g29515; PIDN:CAA34488.1; PID:g29516  
 A:Note: alternative splice form 1  
 R:Vikkula, M.; Metsaeranta, M.; Syvaenen, A.C.; Ala-Kokko, L.; Vuorio, E.; Peltonen, L.  
 Biochem. J. 285, 287-294, 1992  
 A:Title: Structural analysis of the regulatory elements of the type-II procollagen gene.  
 A:Reference number: S24270; MUID:92344585; PMID:1637314  
 A:Accession: S24270  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-28 <VIK>  
 A:Cross-references: EMBL:X58709; GB:S40537; NID:g35659  
 A:Note: this translation is not annotated in GenBank entry HSROCOE1, release 111.0  
 R:Nunez, A.M.; Kohno, K.; Martin, G.R.; Yamada, Y.  
 Gene 44, 11-16, 1986  
 A:Title: Promoter region of the human pro-alpha-1(II)-collagen gene.  
 A:Reference number: A24828; MUID:87031574; PMID:3021582  
 A:Accession: A24828  
 A:Molecule type: DNA  
 A:Residues: 1-8, 'T', 10-28 <NUN>  
 A:Cross-references: GB:M25698; NID:g180872; PIDN:AAA52051.1; PID:g553237  
 R:Baldwin, C.T.; Reginato, A.M.; Smith, C.; Jimenez, S.A.; Prockop, D.J.  
 Biochem. J. 262, 521-528, 1989  
 A:Title: Structure of cDNA clones coding for human type II procollagen. The alpha-1(II)  
 A:Reference number: S06496; MUID:90026318; PMID:2803268  
 A:Accession: S06496  
 A:Molecule type: mRNA  
 A:Residues: 7-28, 'R', '99-157, 'P', 159-440, 'G', 442-456, 'E', 458-640, 'A', 642-831, 'PA', 834, 'F'  
 A:Cross-references: EMBL:X16711; NID:g30040; PIDN:CAA34683.1; PID:g30041  
 A:Note: alternative splice form 1  
 R:Ryan, M.C.; Sandell, L.J.

J. Biol. Chem. 265, 10334-10339, 1990  
 A:Title: Differential expression of a cysteine-rich domain in the amino-terminal prop  
 A:Reference number: A35428; MUID:90285153; PMID:2355003  
 A:Accession: A35428  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 27-81, 'L', 83-103 <RYA2>  
 A:Note: alternative splice form 2: splicing appears to be under developmental regulat  
 R:Su, M.W.; Benson-Chanda, V.; Vissing, H.; Ramirez, F.  
 Genomics 4, 438-441, 1989  
 A:Title: Organization of the exons coding for Pro alpha-1(II) collagen N-propeptide c  
 A:Reference number: A30147; MUID:89233138; PMID:2714801  
 A:Accession: A30147  
 A:Molecule type: DNA  
 A:Residues: 104-157, 'P', 159-236 <SUM>  
 A:Cross-references: GB:J03065; GB:M23660; GB:M25655; GB:M25656; GB:M25730; GB:M32168;  
 R:Ala-Kokko, L.; Baldwin, C.T.; Moskowitz, R.W.; Prockop, D.J.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 6565-6568, 1990  
 A:Title: Single base mutation in the type II procollagen gene (COL2A1) as a cause of  
 A:Reference number: A94227; MUID:90370826; PMID:1975693  
 A:Accession: A33116  
 A:Molecule type: DNA  
 A:Residues: 171-172, 'C', 174-175 <ALA>  
 A:Note: mutant sequence from a family with family with primary generalized osteoarthr  
 R:Diab, M.; Wu, J.J.; Eyre, D.R.  
 Biochem. J. 314, 327-332, 1996  
 A:Title: Collagen type IX from human cartilage: a structural profile of intermolecula  
 A:Reference number: S64673; MUID:96195147; PMID:8660302  
 A:Accession: S64674  
 A:Molecule type: protein  
 A:Residues: 188-189, 'X', 191-195, 1224-1230, 'X', 1232-1236 <DIA>  
 R:Franc, S.; Marzin, E.; Boutillon, M.M.; Lafont, R.; Lechene de la Porte, P.; Herbag  
 Eur. J. Biochem. 234, 125-131, 1995  
 A:Title: Immunohistochemical and biochemical analyses of 20000-25000-year-old fossil  
 A:Reference number: S63514; MUID:96096730; PMID:8529631  
 A:Accession: S63514  
 A:Molecule type: protein  
 A:Residues: 243-261, 575-590, 756-763, 'X', 765-779 <FRA>  
 R:Filler, G.E.; Weiss, M.A.; Colombo, P.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre  
 Am. J. Hum. Genet. 56, 388-395, 1995  
 A:Title: An RNA-splicing mutation (G+51VS20) in the type II collagen gene (COL2A1) in  
 A:Reference number: I38867; MUID:95150028; PMID:7847372  
 A:Accession: I38867  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 440, 'G', 442-456, 'E', 458-480, 'P', 482-509 <TIL1>  
 A:Cross-references: EMBL:U15195; NID:g557053; PIDN:AAB60370.1; PID:g557054  
 R:Ramirez, F.  
 submitted to the EMBL Data Library, December 1988  
 A:Reference number: S04892  
 A:Accession: S04892  
 A:Molecule type: mRNA  
 A:Residues: 501-676, 'A', 678-783, 'A', 785-831, 'PA', 834, 'F', 836-1214 <RAM>  
 A:Cross-references: EMBL:X13783; NID:g30037; PIDN:CAA32030.1; PID:g930050  
 R:Vikkula, M.; Peltonen, L.  
 FEBS Lett. 250, 171-174, 1989  
 A:Title: Structural analyses of the polymorphic area in type II collagen gene.  
 A:Reference number: S05000; MUID:89325561; PMID:2753125  
 A:Accession: S05000  
 A:Molecule type: DNA  
 A:Residues: 630-640, 'A', 642-785 <VIK2>  
 A:Cross-references: EMBL:X16158; NID:g29951; PIDN:CAA34278.1; PID:g1335018; PIDN:CAA3  
 PIDN:CAA34283.1; PID:g1335023; PIDN:CAA34284.1; PID:g1335024  
 R:Bogaert, R.; Tiller, G.E.; Weiss, M.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre  
 J. Biol. Chem. 267, 22522-22526, 1992  
 A:Title: An amino acid substitution (Gly853-->Glu) in the collagen alpha 1(II) chain  
 A:Reference number: A44309; MUID:93054548; PMID:1429602  
 A:Accession: A44309  
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
 A:Molecule type: DNA; mRNA  
 A:Residues: 752-831, 'PA', 834, 'F', 836-1005, 'K', 1007-1036, 'Q', 1038-1052, 'E', 1054-1068, 'E'  
 A:Cross-references: GB:I00977; NID:g180812; PIDN:AAB23914.1; PID:g258774  
 A:Note: sequence extracted from NCBI backbone (NCBIP:117273); parts of this sequence





```

Db      214  ProAlaGlyAlaProGlyProGlnGlyPheGlnGlyAsnProGlyGluProGlyGlu--- 232
QY      1742  CCGGGCCGGAACACCCCTGG---CCTGGTGGGCTCACCACCCACACACA---CCTAGC 1690
Db      233  ProGlyValSerGlyProMetGlyProArgGlyProProGlyProProGlyLysProGly 252
QY      1689  GAGACATCAGAGCAGAGCCCGCAGAGCGCGGTGGAGGTGGAGCGGACGCGCCACTGCT 1630
Db      253  AsnAspGlyGluAlaGlyLysProGlyLysAlaGlyGluArgGly---Pro 268
QY      1629  CCAGCACCCAGCTGTCCATTAGG---AAGGGAGCTCCAGGCTTAGGGCTGGCAGGAAG 1573
Db      269  ProGlyProGlnGlyAlaArgGlyPheProGlyThrProGlyLeu---ProGlyValLys 287
QY      1572  CTGCTCATCAGGCTGCTCTCACTGCTAGCACCCTCAGTGTCCCTCGGTATTTGGGCAGG 1513
Db      288  --GlyHisArgGly--- 291
QY      1512  AACACCTGCTTCTCCGGTGTAGAGGA---GGCCAGTGTGTAGGCGAGGATCTGC 1459
Db      292  -----TyrProGlyLeuAspGlyAlaLysGlyGluAlaGlyAlaProGlyValL 308
QY      1458  AGGCTGAGAGGTGAACCCGGTGAAGCGGCTGAAGCTGTCTCACCAGCGCCACACTGTGG 1399
Db      308  ysGlyGluSerGlySerProGlyGluAsnGly---SerProG 321
QY      1398  GACAGGATGTGGCAGCGGAGCCACAGGGAAGCTGCCACACTGGCCAAATAGACTGCT 1339
Db      321  lyProMetGlyProArgGlyLeuProGlyGluArgGlyArgThrGlyProAlaGlyAla 341
QY      1338  CGAGTGGCGAATCGCT---GCACCAGCGGTGCTCATGACACAGAGAGAAGAC 1292
Db      341  laGlyAlaArgGlyAsnAspGlyGlnProGlyProAlaGlyPro---P 356
QY      1291  CAGGAGATGGCGCACTCAGGAACAGCC---CCAGGCTGCCATCCGACAGCCCTTC 1238
Db      356  roGlyProValGlyProAlaGlyGlyProGlyPheProGlyAlaProGlyAlaLysGly 376
QY      1237  ATCATAGTGTCTCGGGCTCGGTGCGGCTCAGCTCTGGGACGCGCTGTGTACAGCCC 1178
Db      376  luAlaGlyProThrGlyAlaArgGlyProGlyGluAlaGlnGlyProArgGlyGluPro 396
QY      1177  CTCGCCCAAGAAATCCGTGTAACACAGCGTCAAGTCAATGCTCAGCTCAGTCACAG 1118
Db      396  lyThrProGlySerPro---GlyProAlaGlyAlaSerG 408
QY      1117  CTCAGCCACGAAGACCGCGCAGGTCGGGGCATCGGCAGCAGCAGCTGTGTGACCGG 1058
Db      408  lyAsnProGlyThrAspGlyIlePro---GlyAlaLysGlySer---AlaGlyAlaPro 426
QY      1057  GGAAGCAGGCGCCAGGTTCCGGAAGCAACGCGCGCGCATGACAGCAGTGGGG 998
Db      426  lyIleAlaGlyAlaProGlyPheProGlyProArgGlyPro---ProAspProGlnGly 445
QY      997  CGACA----- 993
Db      445  laThrGlyProLeuGlyProLysGlyGlnThrGlyLysProGlyIleAlaGlyPheLys 465
QY      992  --AGGAGGGCCGACAGCCCTTCTGTGGTTCGGTGGGCGCCAGCGCTGCTCTCT 939
Db      465  lyGluGlnGlyProLys-----GlyGluProGlyProAlaGlyProGlnGly 481
QY      938  -----CAGCCACCCAGCTGT 923
Db      481  laProGlyProAlaGlyGluGlyLysArgGlyAlaArgGlyGluProGlyValG 501
QY      922  GGCTCTACGAGGTGAGAAATGAGGTGAGCAGGCGCAAGAGGC----- 876
Db      501  lyProIleGlyProProGlyGluArgGlyAlaProGlyAsnArgGlyPheProGlyGln 521
QY      875  -----ACTCTCTCTCGGTGCCAGGT---AGGGGCCAGGCGACTGTGTCCCA 830

```

RESULT 21

S28052

sucrose transport protein - spinach



US-09-759-143-110 (1-3410) x S43142 (1-533)

241	erLeuThrValLeuAlaLeuSerTyrValIysGluLys-----ProTyrSerP	257
966	CACGACGAAGGGTG-----TCGGCCCTCTCTGTGCGCCC	1001
257	roAspGlnAlaValAspAsnAlaGluAspThrAlaSerGlnAlaSerSerAlaG	277
1002	ACTGCTGTCCATCGCGGGCCGCTTG-----GCTTCCGGAACCTGGCGCCTGCTTC	1055
277	lnProMetProPheGlyGluIleLeuGlyAlaPheLysAsnLeu-----	292
1056	CCCGGCTGCACCAAGCTGTGCTGCCGATGCCCGCACCTCGCGCGGCTTCCTCGGGCTG	1115
293	-----LysArgProMetTrpIleLeuLeuLeuValTr	303
1116	AGCTGTCCAGCTGGATGGCACTCATCACTCTACAGCTGTTTACACGGATTCCTGGGGG	1175
303	hrCysLeuAsnTrpIleAlaTrpPheProPheLeuLeuPheAspThrAspTrpMetGlyA	323
1176	AGGGCTGTATACCAAGGCGTCCCGAGAGCTGAGCCGGGACCCGAGGCGCGAGACACTATG	1235
323	rgGluValTyrGlyGly-----AspSerSerGlySerAlaGluGlnLeuLysLeuTyrA	341
1236	ATGAAGCGCTTCGGATGGGAGCGCTGGGCTGTCTCTGCAGTCGCGCATCTCCCTGGTCT	1295
341	spArgGlyValArgAlaGlyAlaLeuGlyLeuMetLeuAsnSerValValLeuGlyPheT	361
1296	TCTCTCTGGTCATGACCGCGCTGGTGGCGGATTCGGCACTTCGAGCAGTCTATTTGGCCA	1355
361	hrSerLeuGlyValGluValLeuAlaArgGlyValGly---GlyValLysArgLeuTrpG	380
1356	GTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCACATGCTTCCTCCACAGTGTGGCGTGG	1415
380	lyIleValAsnPheValLeu-----AlaValCysLeuAlaMetThrVal---LeuV	396
1416	TGACAGCTTCAGCGCGCCCTCAGCGGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACA	1475
396	alThrLysGlnAlaGluSerThr-----	403
1476	CACTGGCTCCCTTACTACCGGGGAGAGCAGGTGTTCTGGCCCAATACCGAGGGACA	1535
404	-----ArgArgPheAlaThrValS	410
1536	CTGGAGTGCTAGCAGTCAGGACAGCCTGATGACCACTCTCCACGAGCCCTAAGCCTG	1595
410	erGlyGlyAlaLys-----ValPro-----	416
1596	GAGCTCCCTTCCCTAATGACACGCTGGTGCTGAGGACGTGGCTCTCCACCTCCAC	1655
417	-----LeuProProPro-----	420
1656	CGCGCTCTGGGGGCCCTCTCCCTGTGATGTCTCCGTACGTGTGGTGGTGAGAGCCCA	1715
421	-----SerGlyValLysAlaGlyAlaLeuAlaLeuPheAlaValMetGlyValProG	438
1716	CCGAGGCCAGGGTG---GTTCCG-----GGCC	1739
438	lnAlaIleThrTyrSerIleProPheAlaLeuAlaSerIlePheSerAsnThrSerGlyA	458
1740	GGGGCATCTGCCTGACCTCGCCATCTCGGATAGTGCCTCTCTCTCCACGAGTGGCCC	1799
458	laGlyGlnGlyLeuSerLeuGlyValLeuAsnLeuSerIleValIleProGlnMetIleV	478
1800	CATCCCTGTTATGGCTCCATTGTCCAGCTC-----AGCCAGTGTCTCACTGCCTATA	1853
478	alSerValAlaAlaGlyProTrpAspAlaLeuPheGlyGlyGlyAsnLeuProAlaPheV	498
1854	TGGTCTCTCGCGCAGCGCTGGTCTGTGGCCATTACTTCTTGTACAGAGTATATTG	1913
498	alValGlyAlaValAlaAlaLeuAlaSerGlyIlePheAlaLeuThrMetLeuProSerP	518
1914	ACAAGAGCGACTTGGCCAAATCTACGCG	1942
518	roGlnProAspMetProSerAlaLysAla	527



```
QY 2915 CAAGAAGAGAGGGTGGTAGGAAGCGCTTGAGA-----CCTGAAGCCCAACCC 2865
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 ProGlyGluValGlyPheAlaGlySerProGlyAlaArgGlyPheProGlyAlaPro-GI 294
QY 2864 TCTACCTCTCTCAACACCCCTAACCTTGGGTACACAGATTGGAAATATCATATTGGGATG 2805
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 yLeuPro-----GlyLeuLysGlyHisArgGlyHisLysGly-- 306
QY 2804 AGTAGAATTTCCAAGCTCTGGCTAGGCATTTTGGG----- 2768
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 -LeuGluGlyProLysGlyGluValGlyAlaProGlySerLysGlyGlyAlaGlyProTh 326
QY 2767 -----GGGCCA---GACCCAGGAGAGAAGATTCTTGGCAATGATC 2730
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 rGlyProMetGlyAlaMetGlyProLeuGlyProArgGlyMet----- 340
QY 2729 AGCCCAATGACAGCTATCTCAGGACCTGATTGTGGGATCCCCACCCCTACCCAA 2670
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 -----ProGlyGluArgGlyArgLeu-----GlyProGlnGlyAlaProGI 354
QY 2669 TATTAGACACCAACACAGA----- 2649
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 yGlnArgGlyAlaHisGlyMetProGlyLysProGlyProMetGlyProLeuGlyLePr 374
QY 2648 AGCTAGCAATGGATTCCCTCTACTTTGTAAATAAATAAGTTAAATATTAAATGCCTG 2589
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 oGlySerSerGlyPheProGlyAsn----- 382
QY 2588 TGCTCTGTGATGGCAACAGAGAACAGACAGCCACATCTGATAAAGTAAAGGGG 2529
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 -ProGlyMetLysGlyGluAlaGlyProThGly-----AlaArgGlyProGluGI 399
QY 2528 GTGGATCAGAAAAGACAGTCTGTGGCTGAGGGACCTGGTCTTGTGTGTGCC 2469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 yProGlnGlyGlnArgGlyGluThrGlyProGlyProValGlySerProGlyLeuPr 419
QY 2468 CTCGAGCTCTTCCCTACAAATAAGTCATATGTTCAATCCCATCGAGGAGTGTTCAT 2409
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 419 oGlyAlaIleGly----- 423
QY 2408 CTTAGAACTCCCATCAAGAGCTACATTAACAGAACTGCAGTTAAGGGCTTAGAGA 2349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 -----Thras 425
QY 2348 TGGGAACACAGTGAAGTATTTCAGCTCCCAAAACCCCTCTCTAGTGTGTCTCA 2289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 pGlyThrProGly-----ProLysGlyProThr----- 434
QY 2288 ACTAGAGGCTAGCTGTTAACCTGAGCTGGGTAAATCCACCTGCAGAGTCCCGCATTC 2229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 435 -----GlySerProGlyThrSerGlyProProGI 444
QY 2228 CAGTGCATGAGCCCTCTGTGGCTCCCTGTATAGTCCAGACTGAACCCCTTGAAGG 2169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 444 ySerAla---GlyProProGlySerProGly---ProGlnGlySerThrGlyProGlnGI 462
QY 2168 CTTCCAGTACGGAGCCCTAGACACTGGGGAGAGAGG-----AGAGGGAC 2124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 462 y---AsnSerGlyLeuProGlyAspProGlyPheLysGlyGlyAlaGlyProLysGlyGI 481
QY 2123 GCCCAGCCCGGAGCTGTGAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGGCC 2064
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 uProGlyProHisGlyIleGln----- 488
QY 2063 ACATTACTTTGGCAGCAACAGAAACTGGCGCCAGCCCGCCAGCCCATGGGCTAACAG 2004
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 489 -----GlyProIleGlyProProGlyGluGluGlyLy 499
QY 2003 GACGGGGAGCTGGGA---CCCAGTGAGGAGCCCTCCACCCCAATGCTGGAAGTTT 1947
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 sArgGlyProArgGlyAspProGlyThrLeuGlyProPro----- 512
```

---

```
QY 1946 TCTACGCTGAGTATTTGGCCAAAGTCGCTCTTGTGTAATACTACTACTGTGTAGCAAAATAAA 1887
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 -----GlyProVal----- 515
QY 1886 TGGCGACACAGACCCAGCCCTGCGGCAGACACACCATATATAGGCAGTGCACAGACTG- 1835
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 516 -GlyGluArgGlyAlaProGlyAsnArgGlyPheProGlySerAspGlyLeuProGlyPr 535
QY 1834 -----GCTGAGCTGGACCAATGGACCCATAAACAGGATGGGGCCACCTGGGACAGCA 1782
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 oLysGlyAlaGlnGlyGluArgGly-ProValGly-SerSerGlyProLysGlySerGln 554
QY 1781 GGAAGGCACTATCCAGGATGCGAGGTCCAGGCAGATGCCCGCCCGGAACACCCCTGG 1722
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 555 Gly-----AspProGlyArgProGlyGluPro-GlyLeuProGI 567
QY 1721 CTCCTGGTGGCTCACCACACACAGCTAGCGGAGACATCACAGGAGAGA---GCCCGCGC 1665
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 567 yAlaArgGlyLeuThrGlyAsnProGlyValGlnGlyProGlnGlyLysLeuGlyProLe 587
QY 1664 AGAGCGCGGTGGAGGTGGGAGGACCTGCTCCAGCACCCACCCACCTGCTCCATTAGGG- 1606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 587 u-GlyAlaProGlyGluAspGlyArgProGlyProProGly-----SerIleGlyI 604
QY 1605 --AAGGAGCTCCAGGC---TTAGGG---CCTGGCAGGAAGCTGGTTCATCAGGCTGTCT 1554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 604 leLysGlyGlnProGlyThrMetGlyLeuProGlyProLys-GlySerAsnGlyAspPro 623
QY 1553 CACTGTAGCACCTCCAGTGTCCCTCGGTATT-----TGGGAGGAGAACACCTGTCT 1500
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 624 -----GlyLysProGlyGluAlaGlyAsnProGlyVal 634
QY 1499 CCCGGTGGTAGAGGGAGG-----CCAGTGTGTAGGGCAGGA 1464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 635 ProGlyGlnArgGlyAlaProGlyLysAspGlyLysValGlyProTyroGlyProGly 654
QY 1463 TCTGAGGCTGAGAGAGTGAACCGGTGAGGGCGCTGAAGCTGTCCACAGCCACAC 1404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 ProProGlyLeuArg-GlyGluArgGlyGluGlnGly--ProProGlyProThrGlyPhe 673
QY 1403 TGTGGGACAGCATGTGGCAGCCGACCCACAGGAAAG----- 1365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 674 GlnGly-----HisProGlyProProGlyProProGlyGluGlyLysPro 689
QY 1364 -----CTGCCACACTGCCAAATAGACTGCTCGAGTCCCGCAATCGCTGCACCA 1317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 690 GlyAspGlnGlyValProGlyGlyProGlyAlaVal-----GlyPro 703
QY 1316 GCGGTCTCATGACAGAGAGA---AGACCAGGAGATGGCGCACTGCAGGAACAGCCCA 1260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 704 LeuGlyProArgGlyGluArgGlyAsnProGlyGluArgGlyGluProGlyIleThr--- 722
QY 1259 GGCTGCCCATCCGAACGC-----CTTCATCATAGTGTCTCCGGGCGCTCGG- 1215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 723 GlyLeuProGlyGluLysGlyMetAlaGlyGlyHisGlyProAspGlyProLysGlySer 742
QY 1214 -----TGCCCGCTCAGCTGTGGCAGCCCTGGT 1185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 ProGlyProSerGlyThrProGlyAspThrGlyProGlyLeuGlnGlyMetProGly 762
QY 1184 ACA-----GCCCTCGCCACAGAAATCCGTGTAATAACAGCTGAAGGTCA 1140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 763 GluArgGlyIleAlaGlyThrProGlyProLysGlyAsp----- 775
QY 1139 TGAGTGCCATCCAGCTGCACAGCTCAGCCAGAACGCCGCGCAGGTGCGGGCATGC 1080
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 776 -----ArgGlyGlyIleGlyGlyLysGlyAla 784
QY 1079 GGCAGACACACTGGTCGCGGGGAGAGAGCGGCCGCGCAGGTGCGGGCATGC 1020
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 GluGlyThrAlaGlyAsnAspGlyAlaGlyGlyLeuProGly-----ProLeuGly 801
QY 1019 CCCGGCATGCACAGCAGTGGGCGACAAAGAGGGGGCCGACAGCCCTTCTGCTGCTCGG 960
```









1

Db	326	AsnAspGlyValArgAlaGlyAlaLeuGlyLeuMetLeuAsnAlaIleValLeuGlyPhe	345
QY	1295	TTCTCTCTGTGTCATGGACCGCGTGGTGACAGCATTCGGGACTCGACGAGCTCTATTGTGGCC	1354
Db	346	MetSerLeuGlyValGluTrpIleGlyArgLysLeuGly---GlyAlaLysArgLeuTrp	364
QY	1355	ACGTGTGGCAGCTTTTCCTGTGGCTGCGCGGTGCCACATGCTCTCCACACAGTGTGGCGGTG	1414
Db	365	GlyIleValAsnPheIleLeu---AlaIleCysLeuAlaMetThrValVal	381
QY	1414	-----	1414
Db	382	ThrLysGlnAlaGluAsnHisArgAspHisGlyGlyAlaLysThrGlyProProGly	401
QY	1415	---GTGACAGCTTTCAGCC-----GCCCTCACCGGG-----TTTCACC	1447
Db	402	AsnValThrAlaGlyAlaLeuThrLeuPheAlaIleLeuGlyIleProGlnAlaIleThr	421
QY	1448	TTCTCAGCCCTGCAGATCTCGCCTACACACTGGCGCTCCTCTWACCACCGGGAGAG---	1504
Db	422	PheSer-----IleProPheAlaLeuAlaSerIlePheSerThrAsnSerGly	437
QY	1505	-----CAGGTGTCTCTGCCCAANTACCGA	1528
Db	438	AlaGlyGlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValValProGlnMetVal	457
QY	1529	GGGACACTGGAGGTCTAGCAGTGAGGACAGCCTGATCACCAGCTCTCTGCCAGGCCCT	1588
Db	458	IleSerValGlyGlyGlyProPheAspGluLeuPheGlyGlyClyAsnIleProAla---	476
QY	1589	AGCCTGGAGCTCCCTTCCTATGGACACAGTGGTGTGGAGGAGTGGCTGCTC	1645
Db	477	-----PheValLeuGlyAlaIleAlaAlaValSerGlyValLeuAla	491
QY	1646	-----CCACTCCACCCGGG	1660
Db	492	LeuThrValLeuProSerProProProAspAla	502
RESULT	27		
I49607			
procollagen type V alpha 2 - mouse			
C:Species: Mus musculus (house mouse)			
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999			
C:Accession: I49607			
R:Andrikopoulos, K.; Suzuki, H.R.; Solursh, M.; Ramirez, F.			
Dev. Dyn. 195, 113-120, 1992			
A:Title: Localization of pro-alpha 2(V) collagen transcripts in the tissues of			
A:Reference number: I49607; MUID:93214071; PMID:1297453			
A:Accession: I49607			
A>Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: mRNA			
A:Residues: 1-1497 <RES>			
A:Cross-References: GB:L02918; NID:g309180; PIDN:AAA37440.1; PID:g309181			
C:Gene: COL5A2			
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal h			
F:39-98/Domain: von Willebrand factor type C repeat homology <WVC>			
F:1270-1497/Domain: fibrillar collagen carboxyl-terminal homology <FCC>			
Alignment Scores:			
Pred. No.:	5,79e-12	Length:	1497
Score:	323.00	Matches:	333
Percent Similarity:	30.43%	Conservative:	59
Best Local Similarity:	25.85%	Mismatches:	442
Query Match:	5.19%	Indels:	455
DB:	2	Gaps:	74
US-09-759-143-110 (1-3410) x I49607 (1-1497)			
QY	3220	TGCNAACGGCACTTAAACCCCGCTTGAGAGATAGACCTCCCTTAGCTCAGGCAGGGGT	3161
Db	81	CysAlaAsnProIleThrProGlyGlyCysProValCysProGlyCysProGlyGly	100

QY	3160	GCTCCTGAGTTTCTGTGTGATTCGCCCAACACAGCATATACTCTGGGGCTCGAGATGGA	3101
Dd	101	GlyAsp-----ThrSerPheGlyArgGlyArgLysGly	111
QY	3100	CAAAGGCTTGGAACCCGACTTTGTCTTGTGCTCTGAGTAGCTCCAACAGGGTTG	3041
Dd	112	GlnLys---GlyGluProGlyLeuValProValVal-----ThrGlyIle	125
QY	3040	TGGACTGCTGGGAAAGTTGGGG--GTAGGGAAAGTTGGGGTAGGGAAATTITGG	2984
Dd	126	ArgGlyArgProGlyProAlaGlyProProGlySerGln-GlyProArgGlyAspArgG1	145
QY	2983	GCAGTGCCTTCATCACCCAGCTCCTACAGAGAGTAGAGGGAGTGGAAAGTGGGGGAACC	2924
Dd	145	yPro--LysGlyArgProGlyProAArgGlyProGlnGly-----IleAspGlyGluPr	162
QY	2923	AGGCTGGGCCAAGAAGAGGGGTGTAGGGAAGCCGTGTGAGACCTGAAGCCCACCCCT	2864
Dd	162	oglyMet-ProGlyGlnProGlyAlaProGlyProProGlyHisProSerHisProGlyP	182
QY	2863	CTACCTCTCTCAACACCCT-----ValGlyLeuMet-----AAC	2840
Dd	182	roAspLyMetSerArgProPheSerAlaGlnMetAlaGlyLeuAspGluLysSerGlyL	202
QY	2839	TTGGGTAAACAGATTGTGAATTAATCATTTGGGATCAGTAGAATTTCCAAGTCTCTGGG--	2782
Dd	202	euglySerGln--ValGlyLeuMet-----ProGlySe	212
QY	2781	-----TTAGGCATTTTTGGGGGCCAG-----	2761
Dd	212	rValGlyProValGlyProAArgGlyProValGlyLeuGlnGlnGlnGlyAlaG1	232
QY	2760	-----ACCCAGAGAGAA--GATTCCTGGCAATGATCAGCCCCAATGACAGCT	2714
Dd	232	yProAlaGlyProProGlyGluProGlyGluProGly-PrometGlyProIleGlySer-	251
QY	2713	ATCTCAGGGAGCTGATTGTTGGGATCCCCACCCTACCCAAAATATTAGACACCAAAC	2654
Dd	252	-----ArgGlyProGlu-----GlyProProGlyLysProGlyGlu-	263
QY	2653	AGAAAAGCTAGCAATGGATCCCTTCTACTTTGTTAAATAAATAAGTTAAATATAAT	2594
Dd	263	-----	263
QY	2593	GCCTGTCTCTGTATGTCACACAGAGACCACAGCCCATCTCATAAAAGGTAAG	2534
Dd	264	-----AspGly-----GluProGlyArgAsnGlyA	272
QY	2533	AGGGGGTGTGATCAGCAAAAAGACAGTGTGTGGCTGAGGGGA---CCTGGTCTTTGTG	2477
Dd	272	snThrGlyGluValGlyPheSerGlySerProGlyAlaArgGlyPheProGlyAlaProG	292
QY	2476	TGTTGCCCTCAGGACTTCTCCCTFACAATAAAGTCATATGTTCAATFCCCATGGAGAG	2417
Dd	292	lyLeuPro-----GlyLeuLysGlyHisArgGlyH	302
QY	2416	TGTTTCATCTAGAACTCCCATGCAAGAGCTTACATTAAACGAAGCTCCAGTTAAGSGG	2357
Dd	302	isLysGly-LeuGluGlyProLysGlyGluIle-----	312
QY	2356	CTTAGAGATGGGAAACAGGTGACTGAGTTTATTCAGCTTCCAAAAACCTTCTCTAGGT	2297
Dd	313	-----GlyAlaProGlyAlaLysGlyGluAlaGlyProThrGlyPrometGlyAla	329
QY	2296	GTGTCTCAACTAGGAGGCTAGCTGTTAACCCCTGAGCCTGGGTAAATCCACCTGCAGAGTCC	2237
Dd	330	MetGlyProLeuGly-----	334
QY	2236	CCGCAATTCAGTAGGAGCCCTTCTGGCCTCCCTGTATATAGTCCAGACTGAACACCCC	2177
Dd	335	-----ProArgGlyMetPro-----GlyGluArgGly	343

667 GlyProGlyProThrGlyPheGlnGlyLeuProGlyProGlyProGlyGlu 686  
 1283 TGGGGCACTGAGGAACAGCCCGAGCTGCCATCGACCGCCCTTCATCATAGTGTCTCC 1224  
 687 GlyGlyLysAlaGlyAsp---GlnGlyValProGlyGlu-----ProGlyAlaVal 702  
 1223 GGGCTCGGTGCGCGCTGCTGGCAGCGCTGCTGGTACAGCCCTGCGCCACGAAT 1164  
 703 GlyProLeuGlyProArgGlyGluArgGlyAsnProGlyGluArgGlyPro----- 720  
 1163 CCGTGTAAACAGCGTCAAGGTATGAGTCCCATCCAGCTGCACAGCTCAGCCACGAAGA 1104  
 721 -----GlyIleThrGlyLeuProGlyGlu 728  
 1103 GCGGGC-----GCAGGGTGGGGCATCGGG-----AGCAGAGCTGT 1065  
 729 LysGlyMetAlaGlyGlyHisGlyProAspGlyProLysGlyAsnProGlyProThrGly 748  
 1064 GCAGCGGGGAACAGCGGCCCGAGT---TCGGGAAGCCCAAGC----- 1023  
 749 ThrIleGlyAspThrGlyProGlyProGlyLeuGlnGlyMetProGlyGluArgGlyIleAla 768  
 1022 -----GGGCGCGCATGGACAGCAGTGGGGCAGCAAGAGGGCGCGCAGCCCTT 972  
 769 GlyThrProGlyProLysGlyAspArgGlyGlyIleGlyGlyLysGlyAlaGlyThr 788  
 971 CTG-----CTGCTCGGTGG-----GCGCCAGCGCTCGCTCCTCAGCCAGCAGCA 927  
 789 AlaGlyAsnAspGlyAlaArgGlyLeuProGlyProLeuGlyProGlyProAlaGly 808  
 926 GTCTGCTCTCAGCAGGTGGAAGATGAGGTGAGCAGGAGGCAAGAGGCACT----- 873  
 809 LeuLeu-----GlyAlaProGlyGlyProGlyProArgGlyLeuValGly 823  
 872 CCTCTGGGTGCCAGGT-----AGGGG----- 849  
 824 ProProGlySerArgGlyAsnProGlySerArgGlyGlyGluAsnGlyProThrGlyAlaVal 843  
 848 -----CCAGGGCAG 840  
 844 GlyPheAlaGlyProGlnGlySerAspGlyGlnProGlyValLysGlyGluProGlyGlu 863  
 839 TGCTGTCCAGTCATAGTCAGCAGGAGGTAGCCAGCAGGAGCCCGCCAGACTGATCATGA 780  
 864 -----ProGlyGlnLysGlyAspAlaGlySerProGly---ProGlnGly----- 877  
 779 AGCATAGACAGAGTAGGCTGGCAGCAGTGGTCTCGGT----- 741  
 878 -----LeuAlaGlySerProGlyProHisGlyProHisGlyValProGlyLeuLysGly 895  
 740 -----CCGGGAAGGTCAGAGCA 720  
 896 GlyArgGlyThrGlnGlyProProGlyAlaThrGlyPheProGlySerAlaGlyArgVal 915  
 719 GGGCTCCAGTGAAGCAGACCTCGCCACAGAGTCCAGCAGGAGCCCGCCAGCCAGGA 660  
 916 GlyProGlyProAlaGlyAlaProGlyProAlaGlyProAlaGly---GluProGly 934  
 659 TGACAGTCCAGCTCCAGGGCGCTGGATCGGGCAGCAGCAGCGCTCTAGCCAGCGCG 600  
 935 LysGluGlyProProGlyLeuArgGlyAspProGlySerHisGlyArgValGly----- 952  
 599 CCGTTGGGATGAAGAAGGCTCAGCAGGATGCCCAAGCAGAGTCCCGATGAAGGGCC 540  
 953 -----AspArgGlyProAlaGlyProProGlySer----- 962  
 539 GCGCGGGCCATAGCTCCAGCCAGTGGTCACTGGCTAGCGTAGGAGCGGACAGAGA 480  
 963 -----ProGlyAspLys 966  
 479 CCAGGCCAGCAGCTGACCAANTGCCAGCAGCATGGTTCATCACTTCCTCTACCCCA 420

Db 967 GlyAspProGlyGluAspGlyGlnProGlyProAspGly-----ProPro 981  
 QY 419 CTTCCAGCAGCAGAGCGGCACATAGTGTGCTGCGGCCCAACACACACCTCCAGGCCAA 360  
 Db 982 GlyProAlaGlyThrThrGlyGlnArgGlyIleVal-----GlyMetProGlyGln 998  
 QY 359 AGG-----TTAGCAGGTTGACCAAGCAGAGCTGGG 330  
 Db 999 ArgGlyValThrGlyMetProGlyLeuProGlyProAlaGlyThrProGlyLysValGly 1018  
 QY 329 CTTTCCGGTGGCGCAGCAGC-----GGCTACCCACCCAGCCCTCT 291  
 Db 1019 ProThrGlyAlaThrGlyAspLysGlyProProGlyProValGlyProGlySerAsn 1038  
 QY 290 GGACCATAGTGGG---CCAGGGGTAGGCTCAGGGGCGCGTTCAGGCATCCAGAACT 234  
 Db 1039 GlyPro-ValGlyGluProGlyProGluGlyProAlaGlyAsnAspGlyThrProGly-- 1057  
 QY 233 GCTTCTCTCGCTCTGCTCCAGAGCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 174  
 Db 1058 ---ArgAspGlyAlaValGlyGluArgGly-----As 1067  
 QY 173 TAGCAATCAGCAGCGCGCCATTTCTGCCAGCCCTTTGGTCCCGCTCAGCTCTCAGCC 114  
 Db 1067 pArgGlyAspProGly-----ProAlaGlyLeuProGlySerGln-GlyA 1082  
 QY 113 CATGCTCAACACCTCTGCTGTGGGCGCAGCTCAGTGGGACACAGCTCTCATCAGATC 54  
 Db 1082 LaProGlyThrProGlyProValGly---AlaProGlyAspAlaGlyGlnArgGlyGlu 1101  
 QY 53 CTGGC---CGAGGCGCGCGCTGTACCCGAGCAGCGCTGCGAGCTGTGTTCCC 1  
 Db 1101 roGlySerArgGlyProValGlyProGlyProGlyArgAlaGlyLysArgGlyLeuPro 1119  
 RESULT 28  
 S22917  
 collagen alpha 5(IV) chain precursor, renal splice form - human  
 N:Alternate names: procollagen alpha 5(IV) chain  
 N:Contains: collagen alpha 5(IV) chain precursor, leukocyte splice form  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence-revision 27-Feb-1997 #text-change 21-Jul-2000  
 A:Accession: S22917; A57079; A37122; I54317; A34850; S18850; I56971; I76598;  
 R:Zhou, J.; Hertz, J.M.; Leinonen, A.; Tryggvason, K.  
 A:Title: Complete amino acid sequence of the human alpha-5(IV) collagen chain and ide  
 n A1port syndrome patient.  
 A:Reference number: S22917; MUID:92316923; PMID:1352287  
 A:Accession: S22917  
 A:Molecule type: mRNA  
 A:Residues: 1-967 <ZHO>  
 A:Cross-references: GB:M90464; NID:g180826; PIDN:AAA52046.1; PID:g553234  
 R:Zhou, J.; Leinonen, A.; Tryggvason, K.  
 J. Biol. Chem. 269, 6608-6614, 1994  
 A:Title: Structure of the human type IV collagen COL4A5 gene.  
 A:Reference number: A54365; MUID:94165049; PMID:8120014  
 A:Accession: A54365  
 A:Molecule type: DNA  
 A:Residues: 1-922 <ZHD>  
 A:Cross-references: GB:U04470; NID:g463378; GB:U04520; NID:g463428; PIDN:AAC27816.1;  
 R:Zhou, J.; Mochizuki, T.; Smeets, H.; Antignac, C.; Laurila, P.; de Paep, A.; Trygg  
 Science 261, 1167-1169, 1993  
 A:Title: Deletion of the paired alpha5(IV) and alpha6(IV) collagen genes in inherited  
 A:Reference number: A57079; MUID:93361972; PMID:8356449  
 A:Accession: A57079  
 A:Molecule type: DNA  
 A:Residues: 1-27 <ZHA>  
 A:Cross-references: GB:237153; NID:g587203; PIDN:CAA85512.1; PID:g587204  
 R:Pihtajaniemi, T.; Pohjola, E.R.; Myers, J.C.  
 J. Biol. Chem. 265, 13758-13766, 1990  
 A:Title: Complete primary structure of the triple-helical region and the carboxyl-ter  
 A:Reference number: A37122; MUID:90337990; PMID:2380186  
 A:Accession: A37122  
 A:Molecule type: mRNA



A:Residues: 84-439, 'GS', 442-624, 'LAIQ', 629-656, 'FR', 669-887, 'R', 889-1264, 1271-1691 <PIH>  
A:Cross-references: GB:J05558; EMBL:M58526; NID:g1314209  
A:Note: submitted to the EMBL Data Library, February 1991  
A:Note: the authors translated the codon GCC for residue 115 as Val  
R:Renieri, A.; Seri, M.; Myers, J.C.; Pihlajaniemi, T.; Massella, L.; Rizzoni, G.; De Ma  
Hum. Mol. Genet. 1, 127-129, 1992  
A:Title: De novo mutation in the COL4A5 gene converting glycine 325 to glutamic acid in  
A:Reference number: I54317; MUID:93244772; PMID:1363780  
A:Accession: I54317  
A:Molecule type: mRNA  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Residues: 313-324, 'E', 326-330 <REN>  
A:Cross-references: GB:S59334; NID:9299946; PIDN:AAD13909.1; PID:g4261609  
R:Hostikka, S.L.; Body, R.L.; Byers, M.G.; Hoehlyae, M.; Shows, T.B.; Tryggvason, K.  
Proc. Natl. Acad. Sci. U.S.A. 87, 1606-1610, 1990  
A:Title: Identification of a distinct type IV collagen alpha chain with restricted kidne  
A:Reference number: A34850; MUID:90160375; PMID:1689491  
A:Accession: A34850  
A:Molecule type: mRNA  
A:Residues: 914-1264, 1271-1691 <HOS>  
A:Cross-references: EMBL:M31115; NID:g180824; PIDN:AAAS2045.1; PID:g180825  
R:Zhou, J.; Hostikka, S.L.; Chow, L.T.; Tryggvason, K.  
Genomics 9, 1-9, 1991  
A:Title: Characterization of the 3' half of the human type IV collagen alpha-5 gene that  
A:Reference number: A37969; MUID:91169491; PMID:2004755  
A:Accession: S18850  
A:Molecule type: DNA  
A:Residues: 924-1264, 1271-1691 <H3>  
A:Cross-references: EMBL:M63455; EMBL:M63456; EMBL:M63457; EMBL:M63458; EMBL:M63459; EMB  
8; EMBL:M63470; EMBL:M63471; EMBL:M63472; EMBL:M63473; NID:g177922; PIDN:AAAS1558.1; PID  
R:Guo, C.; Van Damme, B.; Van Damme-Lombaerts, R.; Van den Berghe, H.; Cassiman, J.J.; M  
Kidney Int. 44, 1316-1321, 1993  
A:Title: Differential splicing of COL4A5 mRNA in kidney and white blood cells: a complex  
A:Reference number: I56971; MUID:94133540; PMID:8401933  
A:Accession: I56971  
A:Molecule type: mRNA  
A:Residues: 1258-1276 <GUO1>  
A:Cross-references: GB:I569168; NID:g545095; PIDN:AAAS60612.1; PID:g545096  
A:Note: kidney splice form  
A:Accession: I76598  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1264-1291, 'TFLGVLACLV' <GUO2>  
A:Cross-references: GB:S69169; NID:g545097; PIDN:AAAS60613.1; PID:g545098  
A:Note: frameshift mutation in patient with Alport syndrome  
R:Myers, J.C.; Jones, T.A.; Pohjola, E.R.; Kadri, A.S.; Goddard, A.D.; Sheer, D.; S  
Am. J. Hum. Genet. 46, 1024-1033, 1990  
A:Title: Molecular cloning of alpha5(IV) collagen and assignment of the gene to the regi  
A:Reference number: A35335; MUID:90252791; PMID:2339699  
A:Accession: A35335  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1448-1477 <MYE>  
R:Nakazato, H.; Hattori, S.; Ushijima, T.; Matsuura, T.; Koitabashi, Y.; Takada, T.; Yos  
Kidney Int. 46, 1307-1314, 1994  
A:Title: Mutations in the COL4A5 gene in Alport syndrome: a possible mutation in primord  
A:Reference number: I56975; MUID:95156893; PMID:7853788  
A:Accession: I56975  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1595-1602 <NAK>  
A:Cross-references: GB:S75903; NID:g913882; PIDN:AAAS33374.1; PID:g913883  
A:Note: premature termination mutation from a patient with Alport syndrome; one other mu  
R:Lenmink, H.H.; Schroeder, C.H.; Brunner, H.G.; Nelen, M.R.; Zhou, J.; Tryggvason, K.;  
Genomics 17, 485-489, 1993  
A:Title: Identification of four novel mutations in the COL4A5 gene of patients with Alpo  
A:Reference number: I54188; MUID:94010948; PMID:8406498  
A:Accession: I54188  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1604-1607, 'VHDAYKC' <LEM>  
A:Cross-references: GB:S65767; NID:g425563; PIDN:AAD13967.1; PID:g4261667

A:Note: frameshift mutation from a patient with Alport syndrome; five other mutations  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating uni  
ed and subsequently O-glycosylated.  
C:Genetics:  
A:Gene: GDB:COL4A5; ATS  
A:Cross-references: GDB:120596; OMIM:303630  
A:Map position: Xq22-Xq22  
A:Introns: 27/3; 47/3; 77/3; 92/3; 107/3; 128/3; 146/3; 155/3; 182/3; 203/3; 215/3; 215/3;  
3/3; 799/1; 837/1; 893/1; 923/1; 973/1; 1006/1; 1036/1; 1082/3; 1125/1; 1152/1; 1185/1  
A:Note: the alpha 5(IV) and alpha 6(IV) chain genes are encoded on opposite strands w  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha  
mong trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeri  
ng associations in the interrupted helical domain (with disulfide and desmosine cross  
C:Function:  
A:Description: minor structural component of extracellular basement membrane  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: Alport syndrome; basement membrane; coiled coil; extracellular matrix; gl  
F:1-26/Domain: signal sequence #status predicted <Sig>  
F:27-1264/Product: collagen alpha 5(IV) chain, renal splice form #status predicted <M  
F:27-1264,1271-1691/Product: collagen alpha 5(IV) chain, leukocyte splice form #statu  
F:27-41/Domain: amino-terminal nonhelical, NC2 #status predicted <NC2>  
F:42-1462/Region: interrupted helical  
F:1463-1691/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>  
F:1473-1573/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F:1583-1687/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F:29,32,38,40,124,451,481,484/Disulfide bonds: interchain #status predicted  
F:125/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:1482-1570,1535-1573/Disulfide bonds: (or 1482-1573, 1515-1570) #status predicted  
F:1527-1533,1638-1644/Disulfide bonds: #status predicted  
F:1592-1684,1626-1687/Disulfide bonds: (or 1592-1687, 1626-1684) #status predicted  
Alignment Scores:  
Pred. No.: 5,79e-12 Length: 1691  
Score: 323.00 Matches: 284  
Percent Similarity: 30.90% Conservative: 59  
Best Local Similarity: 25.59% Mismatches: 439  
Query Match: 5.03% Indels: 428  
DB: 1 Gaps: 65  
US-09-759-143-110 (1-3410) x S22917 (1-1691)  
QY 2 GGACACGCTGTGACGCGCTGCTCGGTGACAGCCCGCGCCT-----CGG 49  
||||| :||| ||||| |||  
Db 486 GlyThrGlyIleSerGly---ProGlyGlnProGlyLeuProGlyLeuProGlyPro 504  
QY 50 CCAGATCTGAGTGATGAGACGTGCTCCCTGAGTGTGCTGACAGCAGCAGCAGGTTGAG 109  
||||| ||| ||| ||| |||  
Db 505 ProGlySer-----LeuGlyPheProGlyGlnLysGly---Glu 516  
QY 110 CATGGGCTGAGAAGCTGGACCGGACCAAAAGGCTG-----GCAGAAATG 154  
||||| ||||| ||||| |||  
Db 517 LysGlyGlnAlaGlyAlaThrGlyProLysGlyLeuProGlyIleProGlyAlaPro 536  
QY 155 GCGGCTGCTGATCTTCTAGGAGTGTGGCGGACAGCAAGGA----- 194  
||||| ||| ||||| ||| |||  
Db 536 yAlaProGly---PheProGlySerLysGlyGlnProGlyAspIleLeuThrPheProGly 555  
QY 195 -----GGAGAGCGCGCAGCTTCTGAGCAGCAGCAGCGCAGCAGCAGTTC 238  
||||| :||| ||||| |||  
Db 555 yMetLysGlyAspLysGlyGlnLysSerProGlyAlaPro-----GlyLeuPr 572  
QY 239 TGGAGTGTCTGAACGCGCCCTGACGCCCTTACCGCTGCGCCCTGCTGTCAGAGAGGTG 298  
||||| ||| :||| ||||| |||  
Db 572 oGlyLeuProGlyThrProGlyGlnAspGlyLeuProGly-LeuProGlyProLys---- 590  
QY 299 TGGTGTAGCGG-----CCTGCTCGCGCAGCGGAAAGCCAGCTC 337  
||||| ||||| ||||| |||  
Db 591 --GlyLeuProGlyGlyIleThrPheLysGlyGluArgGlyProProGlyAsnProGlyL 610  
QY 338 TTGCTGTCAACCTGCTAACTTTGGCTGTGAGAGTGTGTTTGGCGCGCAGGATCACCTAT 397  
||||| ||| ||||| ||| |||  
Db 610 euProGlyLeuProGlyAsnIleGlyProMetGly----- 621



Db 1138 --ProProGlyAsn-----ProGlyLeuP 1145  
 QY 2477 CACAAGAACAGGTCCTCCAGCCACAGCACTGCTTTTCTGATCCACCCCTCTT 2536  
 |||||  
 Db 1145 roGlyCluProGlyProValGlyGlyHisProGlyGlnProGlyProPro----- 1163  
 QY 2537 ACCTTTATCAGGATGCGCTGTGGCTCTCTGTCCTCCATCAGACAGACAGGCATT 2596  
 |||||  
 Db 1163 ----- 1163  
 QY 2597 TAAATATTTAACTTATTATTATTAACAAGTAGAAGGAATCCATTGCTAGCTTTTCTGTG 2656  
 |||||  
 Db 1163 ----- 1163  
 QY 2657 TTGGTGTCTAAATTTTGGTAGGTTGGGGATCCCAACAATCAGTCCCTGAGATAGC 2716  
 |||||  
 Db 1164 -----GlyGluLysGlyLysProGlyGlnAspGly----- 1173  
 QY 2717 TGGTCATTGGGTGATCATGGCA-----GAATCTTCTCTCTGGGCTGTGG 2764  
 |||||  
 Db 1174 -----ile-ProGlyProAlaGlyGlnLysGlyGluProGly----- 1185  
 QY 2765 CCCCCAAAATGCTAACCCAGGACCTTGGAAATTTCTACTCATCCCAATGATAATCCCA 2824  
 |||||  
 Db 1186 GlnProGlyPheGlyAsnProGlyPro-----Pro 1195  
 QY 2825 AATGCTGTTACCAAGGTTAGGTTGTTGAAGGAGTAGAGGT-----GGGCTTCAG 2878  
 |||||  
 Db 1196 -----GlyLeuProGly--LeuSerGlyGlnLysGlyAspGlyG 1208  
 QY 2879 GTCTCAAGGCTTCCCTACCAACCCCTCTCTCTGGCCAGCGCTGTTCCCTCCCTTC 2938  
 |||||  
 Db 1208 lylLeuProGlylleProGlyAsnProGlyLeuProGlyPro-LysGlyGluProGlyPhe 1227  
 QY 2939 CAC-----TCCCTCTACTCTCTCTAGGACTGGCTGATGAAGGCAC 2980  
 |||||  
 Db 1228 HisGlyPheProGlyValGlnGlyPro----- 1236  
 QY 2981 TGCCCCAAATTTCCCTACCCCACTTTCCCTACCCCACTTTCCCTACCCAGCTCCA 3040  
 |||||  
 Db 1237 -----ProGlyProProGlySerProGlyProAlaLeuGluGlyProLysGly 1252  
 QY 3041 CAACCC 3046  
 |||||  
 Db 1253 AsnPro 1254

## RESULT 29

CGBOIS  
 collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 24-Apr-1984 #sequence\_revision 31-Dec-1993 #text\_change 31-Mar-2000  
 C:Accession: A91193; A91229; A91387; A91211; A91201; A91200; A43048; A02853  
 R:Rauterberg, J.; Timpl, R.; Furthmayr, H.  
 Eur. J. Biochem. 27, 231-237, 1972  
 A:Title: Structural characterization of N-terminal antigenic determinants in calf and hu  
 A:Reference number: A91193; MUID:72255334; PMID:4115172  
 A:Accession: A91193  
 A:Molecule type: protein  
 A:Residues: 1-19 <RAU>  
 A:Experimental source: skin  
 A:Note: the epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is conve  
 R:Fietzek, P.P.; Kuehn, K.  
 Eur. J. Biochem. 52, 77-82, 1975  
 A:Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-bromide  
 A:Reference number: A91229; MUID:76022320; PMID:1164916  
 A:Accession: A91229  
 A:Molecule type: protein  
 A:Residues: 20-145 <FIE>  
 A:Experimental source: skin  
 A:Note: Lys-103 is hydroxylated and binds glucosylgalactose  
 R:Fietzek, P.P.; Wendt, P.; Kell, I.; Kuehn, K.  
 FEBS Lett. 26, 74-76, 1972  
 A:Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from calf

A:Reference number: A91387; MUID:73049499; PMID:4673951  
 A:Accession: A91387  
 A:Molecule type: protein  
 A:Residues: 146-294 <FI2>  
 A:Experimental source: skin  
 R:Fietzek, P.P.; Rexrodt, F.W.; Hopper, K.E.; Kuehn, K.  
 Eur. J. Biochem. 38, 396-400, 1973  
 A:Title: The covalent structure of collagen. 2. The amino-acid sequence of alpha1-CB7  
 A:Reference number: A91211; MUID:74086118; PMID:4359390  
 A:Accession: A91211  
 A:Molecule type: protein  
 A:Residues: 295-562 <FI3>  
 A:Experimental source: skin  
 R:Wendt, P.; Mark, K.V.D.; Rexrodt, F.; Kuehn, K.  
 Eur. J. Biochem. 30, 169-183, 1972  
 A:Title: The covalent structure of collagen. The amino-acid sequence of the 112 resid  
 A:Reference number: A91201; MUID:73042276; PMID:4343808  
 A:Accession: A91201  
 A:Molecule type: protein  
 A:Residues: 563-675 <WEN>  
 A:Experimental source: skin  
 R:Fietzek, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.  
 Eur. J. Biochem. 30, 163-168, 1972  
 A:Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB  
 A:Reference number: A91200; MUID:73042275; PMID:4343807  
 A:Accession: A91200  
 A:Molecule type: protein  
 A:Residues: 676-758 <FI4>  
 A:Experimental source: skin  
 A:Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in pos  
 R:Rauterberg, J.; Fietzek, P.; Rexrodt, F.; Becker, U.; Stark, M.; Kuehn, K.  
 FEBS Lett. 21, 75-79, 1972  
 A:Title: The amino acid sequence of the carboxyterminal nonhelical cross link region  
 A:Reference number: A43048  
 A:Accession: A43048  
 A:Molecule type: protein  
 A:Residues: 759-779 <RA2>  
 A:Experimental source: skin  
 C:Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hyd  
 C:Comment: Prolines in the third position of the tripeptide repeating unit (G-X-Y) ar  
 C:Comment: The order of the eight CNBR peptides in the alpha 1(I) chain of bovine ski  
 9, 149, 268, and 217 residues.  
 C:Comment: The complete chain contains 1052 residues.  
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo  
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trime  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

## Alignment Scores:

Pred. No.:	7.65e-12	Length:	779
Score:	321.00	Matches:	229
Percent Similarity:	31.12%	Conservative:	38
Best Local Similarity:	26.69%	Mismatches:	269
Query Match:	5.16%	Indels:	324
DB:	1	Gaps:	51

US-09-759-143-110 (1-3410) x CGB01S (1-779)

QY	2365	GTTAAGGGCTTAGA-----GATGGAAACACAGGTGACTGAGTTATTCA	2321
Db	102	MeLLysgLyHisArgGlyPheSerGlyLeuAspGlyAlaLysGlyAsp-----Ala	118
QY	2320	GTCTCCAAAACAGCTTCTCTAGGTGTGTCTCAACTAGGAGGCTGTGTTAACCCCTGAGC	2261
Db	119	GlyProAlaGlyProLys-----	124
QY	2260	CTGGGTAAATCCA-----CCTGCAGAGTCCCGCATTCAGTCATGGAGCCCTTCTGCG	2207
Db	125	---GlyGluProGlySerProGlyGluAsnGlyAlaProGlyGlnMetGlyProArgGly	143
QY	2206	CTCCCTGTATAGTCCAGACTGAAACCCCTTTGGAAGGCTCCAGTCAGGACCCCTAGA	2147
Db	144	LeuPro-----GlyPheProGlyProLysGlyAlaLaGlyGluProGly	158



Db	127	AlaGlyGlnGlyProArgGlyAspArgGlyAspGlyGlyGlyGlyGlyGlyAlaProGly	146
Qy	2020	CCCCATGGGGCTAACAGGAGCGGGAGCTGGGA-	1988
Db	147	ProArgGly-----ArgAspGlyGluProGlyThrProGlyAsnProGlyProAlaGly	164
Qy	1987	---CCCAAGTGGAGCGCCCTCCACCCCAATGCTGGAAGTTTCTTACGCTGAGTATTT	1931
Db	165	ProProGlyProProGlyProProGlyLeuSerAlaGlyAsnPhe-----	179
Qy	1930	GGCCAAAGTCGCTTGTCAAATACTACCTGTGTAGCAAAAGTAATGGGACACGCCAG	1871
Db	180	-----AlaAlaGlnMetAlaGlyLysTraspGlu	189
Qy	1870	GCCTGGCGGACACCATATAGCAGCTGACAGACTGGCTGAGCTGGACAATGGAGCCCAT	1811
Db	190	LysAlaGlyGlyAlaGlnMetGly-val-MetGlnGlyProMetGlyProMetGlyPro-	208
Qy	1810	AAACAGGGATGGGGCCACCTGGGACAGCAGGAAGCATTATCCA---GGATGGCGAGT-	1755
Db	209	-----ArgGlyProProGlyProAlaGlyAlaProGlyGlnGlyPheGlnGlyA	226
Qy	1754	-----CCAGCAGATGCCCGCGCGCGGAACCCCTGG---CCTCGGTGGGCTC	1709
Db	226	snProGlyGluProGlyGlu---ProGlyValSerGlyProMetGlyProArg---GlyPr	244
Qy	1708	ACCACCCACACACGTACGGAGACATACAGGCAGAGGCCCGCAGAGCGCGGTGGAG	1649
Db	244	oProGlyProAlaGlyLysProGlyAspAspGlyGluAlaGlyLysProGlyLysSerGl	264
Qy	1648	T-----GGGAGCAGGCCACTGCCTCCAGCCACCCACCTGTCCTCC	1613
Db	264	yGluArgGlyLeuProGlyProGlnGlyAlaArgGlyPheProGlyThrProGlyLeuPr	284
Qy	1612	ATTAGGGAAG-----GGAGCTCCAGGCTTAGGGCTGCGCAGGAAGCTGGTCATCAG	1562
Db	284	oGlyValLysGlyHisArgGlyThrProGlyLeu---AspGlyAlaLys-GlyGluAlaG	303
Qy	1561	GCTGTCCTCACTGCTAGCACCTCCAGTGTCCCTCGGTATTTGGGAGCAGAACCTGCTT	1502
Db	303	lyAlaPro-----GlyVal---LysGlyGluSerGlyS	313
Qy	1501	CTCCCGTGTGTAGAGGAGCGCAGTGTGTAGGCAGGATCTGCAGGCTGAGAAGGTGAA	1442
Db	313	erProGlyGluAsnGlySerPro-----GlyProMetGlyProArg-GlyLeu	328
Qy	1441	CCCGETGAGGCGGCTGAAGCTGTCACCGGGCCACACTGTGGGACAGGCATGTGGCACC	1382
Db	329	ProGlyGluArgGly--ArgThrGlyProAlaGlyAlaAlaGlyAlaArgGlyAsnAspG	348
Qy	1381	GGCAGCCACAGGGAAGCTGCCACACTGGCCAAATAGACTGCTCSAGTCCGGAATCGCTG	1322
Db	348	lyGlnPro-----G	351
Qy	1321	CACGAGCGGTCCATGACAGAGAGACACAGGAGATGGCGCTGTCAGGAACAGCC-	1263
Db	351	lyProAlaGlyPro-----ProGlyProAlaGlyProAlaGlyGlyProG	366
Qy	1262	-----CCAGGCTGCCATCCGAAGCCCTTCATCATAGTGTCTCCGGCCCTCGTGCCCG	1208
Db	366	lyPheProGlyAlaProGlyAlaLysGlyGluAlaGlyProThrGlyAlaArgGlyProG	386
Qy	1207	CTCAGCTGGGCGCCCTGGTACAGCCCTCCGCCACGAAATCGGTAAACACAGCGT	1148
Db	386	lu-----GlyAlaGlnGlySerArgGlyGluProGlyAsnProGlySerProGly-	402
Qy	1147	GAAGGTCAATGATGCCATCCAGCTGCACAGCTCAGCCACGAGAGCGCGCGAGGTGCG	1088
Db	403	-----ProAlaGlyAlaSerGlyAsnProGlyThrAspGlyIlePro---G	417
Qy	1087	GGGCAATGGGACGACACACTGGTCACCGCGGGAAGCAGGCGGCCAGGTTCCGGAAGC	1028

Qy	1648	T	-----GGGAGCAGGCCACTGCCTCCAGCACACCACCGCTGCC	1613
Db	264	yGluArgGlyLeuProGlyProGlnGlyAlaArgGlyPheProGlyThrProGlyLeuPro	284	
Qy	1612	ATTAGGGAAG	-----GGAGCTCCAGGCTTTAGGCCTGGCAGGAAGCTGGTCATCAG	1562
Db	284	oGlyValIysGlyHisArgGlyTyrProGlyLeu	---AspGlyAlaIys-GlyGluAlaG	303
Qy	1561	GCTGTCCTCACTGCTAGCACCTCCAGTGTCCCTCGGTATTGGCGCAGAACACCTGCCT	1502	
Db	303	lyAlaPro	-----GlyVal---LysGlyGluSerGlyS	313
Qy	1501	CTCCCGGTGGTAGGGAGGCCAGTCTGTAGGCAGGATCTGCAGGGCTGAGAAGGTGAA	1442	
Db	313	erProGlyGluAsnGlySerPro	-----GlyProMetGlyProArg-GlyLeu	328
Qy	1441	CCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC	1382	
Db	329	ProGlyGluArgGly--ArgThrGlyProAlaGlyAlaAlaGlyAlaArgGlyAsnAspG	348	
Qy	1381	GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCCGAGTCCGGAATCGCTG	1322	
Db	348	lyGlnPro	-----G	351
Qy	1321	CACACCGCGTTCATGACACAGAGAGAAGACAGGAGATGCGGCTGCGAGGAACAGCC-	1263	
Db	351	lyProAlaGlyPro	-----ProGlyProValGlyProAlaGlyGlyProG	366
Qy	1262	-----CAGGCTGCCATCCGAAGCCTTCATCATAGTGTCTCGGCGCTCGTGCCCGG	1208	
Db	366	lyPheProGlyAlaProGlyAlaIysGlyGluAlaGlyProThrGlyAlaArgGlyProG	386	
Qy	1207	CTCAGCTCTGGGCACGGCCTGGTACAGCCCTCCGCCACGAAATCTGGTGTAAACACAGCGT	1148	
Db	386	lu	-----GlyAlaGlnGlySerArgGlyGluProGlyAsnProGlySerProGly-	402
Qy	1147	GAAGTCAATAGTGCATPCCAGTGCACAGCTCAGCCACGAGAGCGCGCGCAGGTGCG	1088	
Db	403	-----ProAlaGlyAlaSerGlyAsnProGlyThrAspGlyIlePro	---G	417
Qy	1087	GGGCAATCGCGCAGCACACACTGGTCAGCCCGGGAAGCAGGCGCGCCAGGTTCCGGAAGC	1028	

1561	QY	GCTGCTCACTGCTAGCACCTCCAGTGTCCCTCGGTATTTGGCGAGCAACACCTGCCTT	1502
303	Db	lyAlaPro-----GlyVal---LysGlyGlySerGlyS	313
1501	QY	CTCCGGTGTGTAGAGGAGGCGCAGTGTGTAGGCGCAGTCTGCAGCGCTGAGNAAGTGAA	1442
313	Db	erProGlyGluasGlySerPro-----GlyPrometGlyProArg-GlyLeu	328
1441	QY	CCCGGTAGGCGCGCTGAAGTGTCTACCGCGGCACACTGTGTGGCAGGCGCATGTGCACC	1382
329	Db	ProGlyGluArgGly--ArgThrGlyProAlaGlyAlaAlaGlyAlaArgGlyAsnAspG	348
1381	QY	GGCAGCCACAGGGAAGCTGCCACACTGGCGCAANTAGACTGCTCGAGTCCCGAATCGCTG	1322
348	Db	lyGlnPro-----G	351
1321	QY	CACGAGCGGTCCATGACCAGAGAGAAGACCAGGAGATGGCGCACTGCAGGAAACAGCC-	1263
351	Db	lyProAlaGlyPro-----ProGlyProValGlyProAlaGlyGlyProG	366
1262	QY	-----CCAGGTGCTCCATCCGAAGCCTTCATCATGTCTCTCGGCGCTCGGTGGCCCGG	1208
366	Db	lyPheProGlyAlaProGlyAlaLysGlyGluAlaGlyProThrGlyAlaArgGlyProG	386
1207	QY	CTCAGCTTGGCGCGCCCTGGTACAGCCCTCGCCACCAAAATCCGTGTAAACAGCGT	1148
386	Db	lu-----GlyAlaGlnSerArgGlyGluProGlyAsnProGlySerProGly-	402
1147	QY	GAAGGTCAATGATGCCATCCAGCTGCACAGCTACGCCACGAAGAGCGCGCGAGGGTGCG	1088
403	Db	-----ProAlaGlyAlaSerGlyAsnProGlyThrAspGlyIlePro---G	417
1087	QY	GGGCAATGGCGGAGCACACACTGGTGCAGCCGGGGAAGCAGGCGCGCCAGGTTCCGGAAGC	1028

302	lyAlaPro-----	lyGlyGluSerGlyS	313	
303	lyAlaPro-----	GlyVal---	lyGlyGluSerGlyS	313
1501	CTCCGGTGTGTAGAGGAGGCGCAGTGTGTAGGCGCAGGATCTGCAGGGCTGAGNAGTGAA	1442		
313	erProGlyGluAsnGlySerPro-----	GlyPrometGlyProArg-GlyLeu	328	
1441	CCCGGTGAGGCGCGCTGAAGCTGTCACCGGCCACACTGTGGCAGGAGCATGTGCACCC	1382		
329	ProGlyGluArgGly--ArgThrGlyProAlaGlyAlaAlaGlyAlaAspG	348		
1381	GGCAGCCACAGGGAAGCTGCCACACTGGGCCAANTAGACTGCTCGAGTCCCGCAATCGCTG	1322		
348	lyGlnPro-----	lyGlnPro-----	351	
1321	CACGAGCGGTCCATGTACCAGAGAGAACACAGGAGATGGCGCACTGCAGGAAACAGCC-	1263		
351	lyProAlaGlyPro-----	ProGlyProValGlyProAlaGlyGlyProG	366	
1262	-----CCAGGCTGCCATPCGAAGCCCTTCATCATAGTGTCTCGGGCCCTCGGTGGCCCGG	1208		
366	lyPheProGlyAlaProGlyAlaLysGlyGluAlaGlyProThrGlyAlaArgGlyProG	386		
1207	CTCAGCTCGGCAGCCCGCTGGTACAGCCCTCGCCACCAAAATCCGTGTAAACAGCGT	1148		
386	lu-----	GlyAlaGlnSerArgGlyGluProGlyAlaAsnProGlySerProGly-	402	
1147	GAAGGTGATAGTGCCATCCAGCTGTCACAGTCTACGCCACGACGAGCGCGGCGAGGGTGGC	1088		
403	-----	ProAlaGlyAlaAsnProGlyThrAspGlyLeuPro---	417	
1087	GGGCAATGGCGAGCACACTGGTGCAGCCCGGGAAGCAGGCGGCCAGGTTCGGAAAGC	1028		

Db 417 lyAlaLysGlySer---AlaGlyAlaProGlyIleAlaGlyAlaProGlyPheProGlyP 436  
 QY 1027 CAAGCGGGCGCGATGACAGCAGTGGCGGACA----- 993  
 Db 436 roArgGlyPro---ProGlyProGlnGlyAlaThrGlyProLeuGlyProLysGlyGlnA 455  
 QY 992 -----AGGAGGGGGCGCAGAGCCCTTCTGCTGG 965  
 Db 455 laGlyGluProGlyIleAlaGlyPheLysGlyAspGlnGlyProLys-----G 471  
 QY 964 CTGGTGGGGCGCAGGCTCCCTCAGCCAGCAGTGGCTGCTACGCGAGTGA 905  
 Db 471 lyGluThrGlyProAlaGlyProGlnGlyAlaPro-----G 483  
 QY 904 GAGATGAGGGTGAGCAGGCGCAAGAGGCACT----- 873  
 Db 483 lyProAlaGlyGluGlyLysArgGlyAlaArgGlyGluProGlyGlyAlaGlyProI 503  
 QY 873 ----- 873  
 Db 503 leGlyProProGlyGluArgGlyAlaProGlyAsnArgGlyPheProGlyGlnAspGlyL 523  
 QY 872 -----CCTCTGGGTGCCAGGT---AGGGGGCGCAGGCCTGGTCTCCAGTCAAT 824  
 Db 523 euAlaGlyProLysGlyAlaProGlyGluArgGlyProSerGlyLeuAlaGlyProLysG 543  
 QY 823 GCGAGCGAGA-----GGTAGCCAGCAGGCCCCAGCATGATCATGAAGGCATAGAC 770  
 Db 543 lyAlaAsnGlyAspProGlyArgProGlyGluProGlyLeuProGlyAlaArgGlyLeu 563  
 QY 769 AGAGTAGGCTGCGGACAGTGTCCGGTCCCGGAGAGGTGAGAGCAGAGCGGCTCCAG 710  
 Db 563 hreGlyArgProGlyAspAlaGlyPro---GlnGlyLysValGlyProSerGlyAlaProG 582  
 QY 709 TGGAGTGAAGCACACCTGCCACAGATGCCAGAGTCCAGCAGCCCGCCAGC----- 666  
 Db 582 lyGluAspGlyArgProGlyProProGlyProGlnGlyAlaArgGlyGlnProGlyValM 602  
 QY 665 -----CCAGATGACAGTGCAGCTCCAGCTCCAGGGCGCTGGATCCGGGCACAGAGCCC 614  
 Db 602 etGlyPheProGly-----ProLysGlyAlaAsnGlyGluProGlyLysAlaGly- 618  
 QY 613 TGCTAGCCAGCGCGCTTGGATGAGAAAGAGGTCTCAGCAGGATGCCCA-----AGGA 560  
 Db 619 -----GluLysGlyLeuAlaGlyAlaProGlyLeuArg 630  
 QY 559 CAGTGCACAGATGAAGGGCGCGCGGCATAGCGTCCAGCCAGTGGTCACTGGCTGA 500  
 Db 630 lyLeuPro-----GlyLysAspGlyG 637  
 QY 499 GCCTAGGAGCGGACACAGACCAGGCCCGCA-----CTGGACC 461  
 Db 637 luThrGlyAlaAlaGlyProProGlyProSerGlyProAlaGlyGluArgGlyGlnG 657  
 QY 460 AATGCCACGACCATGGTATGAATCTCTCTACCCCTCCAGCAGCAGAGCGCG 401  
 Db 657 lyAlaProGlyPro-----SerGlyPheGlnGlyLeuProGly----- 669  
 QY 400 CACATAGGTGATCGCTCGCGCCCAACACACCTCCAGGC-----CAAAAGGTAG 353  
 Db 670 -----ProProGlyProProGlyGluGlyGlnGlyLysGlnGlyAspG 683  
 QY 352 CAGGTTGACCAAGAGCTGGGCTTTCGGTGGCGCAGCA----- 312  
 Db 683 lnglyleProGlyCluAlaGlyAlaProGlyLeuValGlyProArgGlyGluArgGlyP 703  
 QY 311 -----GGGGGTACCCACAGCCTCTGGACCATAGTGGCGCCAGCGGG---TAG 266  
 Db 703 heProGlyGluArgGlySerProGlyAlaGlnGly-LeuGlnGlyProArgGlyLeuPro 722  
 QY 265 GCCTCAGGGGGCGGTTCAGGCACCTCCAGAACTGCTCTCGGCTCTGCTCCAGAAAGCT 206  
 Db 723 GlyThrProGlyThrAspGly---ProLysGlyAlaAlaGlyProAspGlyProGly 741

QY 205 GCGGCTCTCCCTTGTCTGCGCAACTGCTAGTAATCAGCCAGGCGCCCATTTCTGC 146  
 Db 742 AlaGlnGlyProProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyIleAla 761  
 QY 145 CAGCCCTTTGGT-----GCCGCTCCAGCTTCTCAGCCC----- 113  
 Db 762 GlyProLysGlyAspArgGlyAspValGlyGluLysGlyProGluGlyAlaProGlyLys 781  
 QY 112 -----ATGCTCAACACCTCTGCTGTGGGGCACCTCAGTGG 77  
 Db 782 AspGlyGlyArgGlyLeuThrGlyProIleGly-ProProGlyProAlaGlyAlaAsnG 801  
 QY 76 GCACACGCTCATCATCTCAGATCTGGC-----CGAGCGCGCGCGCTGTGCACC 29  
 Db 801 yGluLysGlyGluValGlyProProGlyProSerGlySerThrGlyAlaArgGlyAlaPr 821  
 QY 28 CGGA 25  
 Db 821 OGly 822  
 RESULT 31  
 S48789  
 sucrose transport protein - common tobacco  
 C:Species: Nicotiana tabacum (common tobacco)  
 C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 26-May-2000  
 C:Accession: S48789  
 R:Buerkle, X.Y.Z.: Frommer, W.B.  
 submitted to the EMBL Data Library, October 1994  
 A:Reference number: S48789  
 A:Accession: S48789  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-507 <BUDE>  
 A:Cross-references: EMBL:X82276; NID:g575350; PID:g575351  
 C:Superfamily: common tobacco sucrose transport protein ;

## Alignment Scores:

Pred. No.: 9,43e-12 Length: 507  
 Score: 319.50 Matches: 126  
 Percent similarity: 41.04% Conservative: 87  
 Best Local Similarity: 24.28% Mismatches: 207  
 Query Match: 4.98% Indels: 99  
 DB: 2 Gaps: 18

US-09-759-143-110 (1-3410) x S48789 (1-507)

QY 278 CCCACTATGTCAGAGGCTGTGGGTGAGCGCGCTGCTGGGCACCGGAAGCCAGCTC 337  
 Db 18 ProLeuProLysSerLysLeuTrp-----Lysile 27  
 QY 338 TTGCTGTCTAACCTTGCTTAACCTTTGGCTGGAGGTGTGTTGGCCGCGCAGCATCACCTAT 397  
 Db 28 IleMetValAlaSerIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeu 47  
 QY 398 GTCCCGCTCTGCTGTGGAGGTGGGGGTAGAGGAGTTCATGACCATGTGCTGGGC 457  
 Db 48 LeuThrProTyrValGlnLeuLeuGlyIleProHisLysPheAlaSerPheIleTrpLeu 67  
 QY 458 ATTGTGTCAGTGTGGCGCTGTGTCCTCCGCTCTAGGCTCAGCCAGTACCACTGG 517  
 Db 68 CysGlyProIleSerGlyMetIleValGlnProValGlyTyrTrpSerAspAsnCys 87  
 QY 518 CGTGGAGCTATGGCGCGCGCGCTTCTATCTGGGCACCTGCTGGCATCTCTGCTGG 577  
 Db 88 SerSerArgPheGlyArgArgGlyPheIleAlaAlaGlyAlaLeuValThrIle 107  
 QY 578 AGCCTCTTCTCATCCCAAGCGCGCTGGCTAGCAGGCTGTGTGCCGATGCC--- 634  
 Db 108 AlaValPheLeuIleGlyPheAlaAlaAspLeu---GlyHisAlaThrGlyAspProLeu 126  
 QY 635 -----AGGCCCTGGAGCTGGCAGCTCATCTGCGCTGGGCTGCTGGAC 682







C:Species: Homo sapiens (man)  
 C:Date: 31-Mar-1990 #sequence\_revision 03-Oct-1995 #text\_change 08-May-1998  
 C:Accession: A35239; A31795

R:Yoshioka, H.; Ramirez, F.  
 J. Biol. Chem. 265, 6423-6426, 1990  
 A:Title: Pro-alpha1(XI) collagen. Structure of the amino-terminal propeptide and expression of the alpha1(XI) chain in fetal rat aorta.  
 A:Reference number: A35239; PMID:90202924; PMID:1690726

A:Accession: A35239  
 A:Molecule type: mRNA  
 A:Residues: 1-558 <YOS>  
 A:Cross-references: GB:J05407  
 R:Bernard, M.; Yoshioka, H.; Rodriguez, E.; van der Rest, M.; Kimura, T.; Ninomiya, Y.; J. Biol. Chem. 263, 17159-17166, 1988  
 A:Title: Cloning and sequencing of pro-alpha1(XI) collagen cDNA demonstrates that type X cartilaginous tissue.

A:Reference number: A92689; PMID:89034222; PMID:3182841  
 A:Accession: A31795  
 A:Molecule type: DNA: mRNA  
 A:Residues: 538-1806 <BER>  
 A:Cross-references: GB:J04177  
 A:Note: parts of this sequence were determined by protein sequencing  
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit (X) and subsequently O-glycosylated.

C:Genetics:  
 A:Gene: GDB:COL1A1; COL16  
 A:Cross-references: GDB:120595; OMIM:120280  
 A:Map position: 1p21-1p21  
 A:Introns: 561/3; 579/3; 597/3; 615/3; 633/3; 648/3; 666/3; 681/3  
 A:Note: the list of introns is incomplete  
 C:Complex: type XI collagen may be a heterotrimer of two alpha 1(XI) chains and one alpha 3(XI) chain (see PIR:CGH06C), initially linked by disulfide bonds among their carboxyl-termini with desmosine cross-links made from lysine and allysine residues

C:Function:  
 A:Description: structural component of extracellular fibrous polymer associated with cell adhesion  
 A:Note: may play a role in controlling the lateral growth of collagen II fibrils  
 C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology  
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproline

F:1-36/Domain: signal sequence #status predicted <SIG>  
 F:35-260/Domain: PAPP-like #status predicted <PAPP>  
 F:37-511/Domain: amino-terminal propeptide #status predicted <PRO>  
 F:512-1565/Product: collagen alpha 1(XI) chain #status predicted <MAT>  
 F:512-527/Region: amino-terminal nonhelical telopeptide  
 F:528-1542/Region: helical  
 F:1543-1565/Region: carboxyl-terminal nonhelical telopeptide  
 F:1566-1806/Domain: carboxyl-terminal propeptide #status predicted <Cnp>  
 F:1583-1805/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
 F:61-243,182-236/Disulfide bonds: #status predicted  
 F:505/Modified site: allysine (Lys) #status predicted  
 F:612,1452/Modified site: 5-hydroxylysine (Lys) #status predicted  
 F:612,1452/Binding site: carbohydrate (Lys) (covalent) #status predicted

Alignment Scores:  
 Pred. No.: 1806  
 Score: 1.54e-11  
 Length: 1806  
 Matches: 315  
 Conservative: 61  
 Percent Similarity: 31.07%  
 Best Local Similarity: 26.03%  
 Query Match: 5.08%  
 Indels: 423  
 Gaps: 71  
 Db:

US-09-759-143-110 (1-3410) x CGHUIE (1-1806)

QY 3101 ACAAGGCTGGGAACCGCACTTTGTCTCTGCT----- 3066

Db 409 ThrAspIleThrGluThrSerIleAsnGlyHisGlyAlaThrGlyGluLysGlyGlnLys 428

QY 3065 -----CCTGCAGTAGCT---CCAAACAGGCTGTGGAGCTGGTGGGAAGTTGGGGGT 3015

Db 429 GlyGluProAlaValGluProGlyMetLeuValGlu-----GlyPro 443

QY 3014 AGGGAAAGTTGGGGTAGGGAATTTTGGGAGTGCCTTCATCAGCCAGCTCCTAGAG 2955

Db 444 ProGlyProAlaGlyProAlaGlyIleMetGly-----ProProGlyLeuGln 459

QY 2954 AGAGTAGAGGGGAGTGGGAAGTGGGGGAACAGGCTGGGCAAGAGAAGAGGGTGTGTTA 2895  
 Db 460 -----GlyProThrGlyProGlyAspProGlyAspGlyProPro 474  
 QY 2894 GGAAGCGCTTGGAGCTGAAGCC-----CCACCTCTACCTCTCTTCAA 2850  
 Db 475 GlyArgProGlyLeuProGlyAlaAspGlyLeuProGlyProGlyThrMetLeuMet 494  
 QY 2849 CACCCTACCTTGGGTAAACACCATTTGGAATATCATTTGGGAGTGAAGATTTCCAAG 2790  
 Db 495 LeuPro-----PheArgTyrGlyGlyAsp----- 502  
 QY 2789 GTCCTGGGTAGGCACTTTGGGGGCCGAGA-----CCCCAGGAGAAGAAG 2745  
 Db 503 -----GlySerLysGlyProThrIleSerAlaGlnGluAlaGln 515  
 QY 2744 ATTCTGCAATGATCAGCCCAATCAGCAGCTATCTCAGGGGACCTGAT----- 2697  
 Db 516 AlaGlnAlaIleLeuGlnGlnAlaArgIleAlaLeuArgGlyProGlyProMetGly 535  
 QY 2696 ---TGTTGGGATCCCCCACCCTACCCCAATATTAGACACCAACAGAAAGCTAGCAA 2640  
 Db 536 LeuThr-GlyArgProGlyProValGlyGlyProGlySerSerGlyAlaLysGlyGluSe 555  
 QY 2639 TGGATTCCTCTTACTTTTAAATAAATAAGTTAAATATTAAATGCTGCTCTCTGT 2580  
 Db 555 rGlyAspProGlyPro-----GlnGlyProArgGlyVa 566  
 QY 2579 GATGCAACAGAGGACCAACAGGCCACATCTGTATAAAG----- 2538  
 Db 566 lGlnGlyProGlyProThrGly--LysProGlyLysArgGlyArgProGlyAlaAsp 585  
 QY 2537 -----TAAGAGGGGTGGATCAGCAAAAGACAGCTGCTGGGCTG 2496  
 Db 586 GlyGlyArgGlyMetProGlyGluProGlyAlaLysGlyAspArgGlyPheAspGlyLeu 605  
 QY 2495 AGGGGA---CCTGGTCTTGTGTGTGCTGCCCTCAGGACTCTTCCCTACAAATAAGTCAT 2439  
 Db 606 ProGlyLeuProGlyAsp----- 611  
 QY 2438 ATGTTCAATCCCTGGAGGAGTGTTCATCTCTAGAACTCCCATGCA----- 2391  
 Db 612 -----LysGlyHisArgGlyGluArgGlyProGlnGlyProGlyProProGlyAsp 629  
 QY 2390 -----AGAGCTACTATAACCAAGCTGCA 2367  
 Db 630 AspGlyMetArgGlyGluAspGlyGluIleGlyProArgGlyLeuProGlyGluAlaGly 649  
 QY 2366 GGTAAAGGGCTTAGAGAT-----GGGAAACAGGTGACTGAGTTTATTCAGCTCCCAA 2313  
 Db 650 -ProArgGlyLeuLeuGlyProArgGlyThrProGly-----AlaProGly 664  
 QY 2312 AAACCTTCTCTA---GGTGTCTCACTAGAGGCTAGTGTTAACCTGAGCCTGGG 2256  
 Db 664 yGlnProGlyMetAlaGlyValAspGlyProGly-----ProLysGly 679  
 QY 2255 TAAT-----CCACCTGCAGAGTCCCGCATTCAGTGCATG 2220  
 Db 679 yAsnMetGlyProGlnGlyGluProGlyProGlyGlnGlyAsnPro----- 696  
 QY 2219 GAGCCCTTCTGCTCCCT---GTATACTCAGACTGAACCCCTTGGAGGCTTCCA 2163  
 Db 697 -GlyProGlnGlyLeuProGlyPro-GlnGlyProIleGlyProGlyGlyGlyPro 716  
 QY 2162 GTCAG-----GCAGCCCTA----- 2149  
 Db 716 roGlnGlyLysProGlyLeuAlaGlyLeuProGlyAlaAspGlyProProGlyHisProG 736  
 QY 2148 -----GAGACTGGGAGAGAGAGAG-----GGAGCCCGCCAGCTGTGCA 2103  
 Db 736 lylLysGluGlnSerGlyGluLysGlyAlaLeuGlyProGlyProGlnGlyProI 756  
 QY 2102 GCTACGACCTCAGCAGCAGGCTGGCAGCAGAGAGCCCATTTACTTTGGCAGCAGACAG 2043

Db	756	leGly***ProGlyProArgGlyValLysGlyAlaAspGlyValArgGlyLeuLysGly	776
QY	2042	AACTGGC-----GGCCAGCCCGCCAGCCCC-----ATGGGGCTAAACAG	2004
Db	776	erLysGlyGluLysGlyGluAspGlyPheProGlyPheLysGlyAspMetGlyLeuLysG	796
QY	2003	GAGC-----CGGAGCTGGCACCC-----	1985
Db	796	lyAspArgGlyGluValGlyGlnIleGlyProArgGly***AspGlyProGluGlyProL	816
QY	1984	--ACTGAGCGAGCCCTCCACCCCAATGCTGCAAGATTTCTACGCTGAGTATTGCGC	1927
Db	816	ysGlyArgAlaGlyProThrGlyAspProGlyProSer-----GlyG	830
QY	1926	AAGTCGCTCTGTCAATACTACTCTGTGTAGCAAAAGTAAATGGCGCACACAGAC	1875
Db	830	lnAla-----GlyLysGlyLysLeuG	838
QY	1874	-----CCAGGCTCGGCAGACACATATAGGCAGTGCACAGACTGGCTGAGCTGACAAT	1820
Db	838	lyValProGlyLeuProGlyTyrProGlyArgGln-----	849
QY	1819	GGAGCCCAATAACAGGATGGGCCCTCGGCACAGCAGGAAGCACTATCCAGATGG-	1761
Db	850	-----GlyProLysGlySerThrGly-----PheProGlyPheP	861
QY	1760	-----CSAGTCCAGGCAGATGCCCGGCCGGAA-----	1731
Db	861	roGlyAlaAsnGlyLysGlyAlaArgGlyValAlaGlyLysProGlyProArgGlyG	881
QY	1730	-----CCACCCTGGCTCGGTGGCTCACCCACCACACAGTACCGAGACATCA	1679
Db	881	lnArgGlyProThrGlyProArg--GlySer--ArgGlyAlaArgGlyProThrGlyL	899
QY	1678	GGCAGAGCCCGCCAGAGCGGGTGGAGTGGGAGCAGGCCACTGCCTCCACGCCACC	1619
Db	899	s----ProGlyProLysGlyThrSerGlyLysGlyAspGlyProGlyProGlyGluAr	918
QY	1618	GTGTCCATTAGGAGGAGGAGCTCCAGGCTTAGGGCCCTGGCAGGAAGCTGTCATCAG	1559
Db	918	gGlyProGlnGlyProGlnGlyProValGlyPheProGlyProLys-----	933
QY	1558	GTCCTCACTGCTAGCACTCCAGTGTCCCTCGTATTTGGCAGGAACACCTGCTCTC	1499
Db	934	-----GlyProProGlyProPro-Gly--ArgMetGlyCysProGlyHisP	948
QY	1498	CCGCTGTGTAGGAGGAGCCAGTGTGTAGGCGAGGATCTGCAGGGCTGAGAAGTGA	1439
Db	948	roGlyGlnArgGly-----GluThrGly-PheGlnGlyLysThrGlyProPro	963
QY	1438	GGTGAGCGCGCTCAAGCTGTCACCAGCGCCACACTGTGGACAGGAGTGTGCACCGG	1379
Db	964	GlyProGlyGly-----ValValGlyProGlnGlyProThrGly	976
QY	1378	AGCCACAGGGAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTCCGAATCGCTGC	1319
Db	977	Glu-ThrGly-----Pro-IleGlyGluA	984
QY	1318	CACCGGTCCATCACCAGAGAGAACAGGAGATGGCGCACTGCAGAACAGCC-----	1263
Db	984	rgGlyTyrProGlyProProGlyProProGlyGluGlnGlyLeuProGlyAlaAlaGlyL	1004
QY	1262	--CCAGGCTGCCATCCGAGGCGCTTCATCATAGTCTCTCCGGCGCTCGGTCCCGG	1205
Db	1004	ysGluGlyAlaLysGlyAspProGlyProGlnGlyIleSerGlyLysAspGlyProAla	1024
QY	1204	AGTCTGGGACGCCCTGTGTACA-----GCCCT-----	1176
Db	1024	lyLeuArgGlyPheProGlyGluArgGlyLeuProGlyAlaGlnGlyAlaProGlyLeuL	1044
QY	1175	-----CGCCACCGAAATCCGTGTAAACACCGTGAAGGTGATGAGTGCATCAGTGC	1127

Db	1044	ysGlyGlyGluGlyProGlnGlyPro-----ProGlyp	1051
QY	1126	GCTGCACAGCTCAGCCACGAAGACCGCGCAGGCTCGGGGCATCGGCAGCACAGCTG	1067
Db	1055	roValGlySerProGlyGluArgGlySerAlaGly-----ThrAlaG	1069
QY	1066	GTGCAGCGCGGAAGCAGCGGCGCCAGGTTCGCGAAGACCAAGCGGCGCCGCGCATGGACA	1007
Db	1069	lyProIleGlyLeuArgGlyArgProGlyProGlnGlyProGlyProAlaGlyGluL	1089
QY	1006	GCAGTGGGGCCACAAGAGAGGGGGCCACAGCCCTTCTGCTGGTGGGGGGCCAGCGC	947
Db	1089	ysGlyAlaProGlyGluLysGlyProGln-----GlyProAla-	1101
QY	946	TGCCTCTCAGCCACCAGCAGTAGTGGGCTGTCTACGAGTGAAGAGATGAGGCTGAGCAG	887
Db	1102	-----GlyArgSpGlyValGlnG	1108
QY	886	GCCAAAGAGGCACCTCCTCGGTGGGCCAGST--AGGGGGCCAGGGCACTGGTGTCCCA	830
Db	1108	ly-----ProValGlyLeuProGlyProAlaGlyPro-----	1118
QY	829	GTCAATGGCAGGAGGAGGTAGCCACAGGCAGCCGCCAAGACTGATCATGAAGCATAGAC	770
Db	1119	-----AlaGlySerProGlyGluAspGlyAsp-----LysGlyGluI	1131
QY	769	AGACTAGGCTGGGCACAGTGGTCCGGGTCCCGAAGAGGTCTACAGCAGCAGGCGCTCCAG	710
Db	1131	leGlyGluProGlyGlnLysGlySerLysGlyLysGlyGluAsn---GlyProPro-	1149
QY	709	TGGAGTGAAGCACACCTGGCCACAGAAGTCCACAGCAGCCGCCAGGATGACGAGTGC	650
Db	1150	-----GlyProProGlyLeuGlnGlyProValGly---AlaProGly-----I	1163
QY	649	CAGCTCCAGGGCGCTGGATCCGGGCACAGCAGCCCTGTCTAGCCAGCCGCGCTTGGGAT	590
Db	1163	leAlaGlyGlyAspGlyGluProGly-----	1171
QY	589	GAGAAAGAGGCTCAGCAGGATGCCAAGGACAGTGCCACAGTGAAGGGCGCGCGCGGC	530
Db	1172	--ProArgGly-GlnGlnGlyMetPheGlyGlnLysGlyAspGluGly---AlaArgGly	1189
QY	529	ATACGCTCCAGCGCAGTGGTCTAGGCTAGGAGCGGGACACACAGCAGGCGCCAG	470
Db	1190	PheProGlyProPro--Gly-----ProIleGlyLeuGlnGlyLeuProGlyPro	1206
QY	469	CACTGGACCAATGCCCA-----GCACCATGTGTCGAACCTTCCTCTCAACCC	422
Db	1206	roGlyGluLysGlyGluAsnGlyAspValGlyProTrpGly-----ProProGlyProp	1224
QY	421	CACCTCCAGCAGCAGAGGGCGGCATAGGTGATGCCTGCGGGCAACACACCTCCAGCGC	362
Db	1224	roGlyProArgGlyProGlnGlyProAsnGlyAlaAspGlyProGlnGlyProProGly-	1243
QY	361	AAAGGTAGCAGGTTGACCAGCAAGAGCTGGGCTTCCGGTCCCGCAGCAGCGCGCTCAC	302
Db	1244	-----SerValGlySerValGlyValGlyGluLysGlyGluP	1257
QY	301	CCACAGCCTCTGGACCATAGTGGG---CCAGCGGGGTAGGGCTCAGGGGGCGGTTCAGGC	245
Db	1257	roGlyGluAlaGly-AsnProGlyProProGlyGlyAlaGlyValGlyGlyProLysGly	1276
QY	244	ACTCCAGAACTGCTTCGCTCGCTCTGCT---CCAGAAGCTGCCGGCTCTCCTCTTGC	188
Db	1277	-----GluArgGlyGluLysGlyGluAlaGlyProProGlyValAlaGlyProProGly	1294
QY	187	TGCCGC-----CAACTGCTAGGAATCAGCCAGGGCGCCCATTTCT-----	148
Db	1295	AlaLysGlyProProGlyAspAspGlyProLysGlyAsnProGly-ProValGlyPhePr	1314
QY	147	-GCCAGCCCTTTGGTCCCGGTCCAGCTTCTCAGCCCATGCTCAACACTGTGCTGTGGG	89
Db	1314	oGlyAspProGlyProGlyGluLeu-----GlyProAlaGlyGlnAs	1329

QY 88 GCACCTCTAGTGGGACACGCTCTCATCTACCTGATCCTGGCCGA----- 46  
Db 1329 pGlyValGlyGlyAspLysGlyGluAspLysProGlyGlnProGlyProGlyPr 1349  
QY 45 -----GGCGCGCGCTGTCTACCCGGAGCCAGC 19  
Db 1349 oSerGlyGluAlaGlyProGlyProGlyLysArgGlyProGlyAlaAla 1368  
RESULT 34  
S16366  
collagen alpha 2(IV) chain precursor - pig roundworm  
C:Species: Ascaris suum (pig roundworm)  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 13-Aug-1999  
C:Accession: S16366  
R:Pettitt, J.; Kingston, I.B.  
J. Biol. Chem. 266, 16149-16156, 1991  
A:Title: The complete primary structure of a nematode alpha-2(IV) collagen and the partial  
A:Reference number: S16366; MUID:91340768; PMID:1714907  
A:Accession: S16366  
A:Molecule type: mRNA  
A:Residues: 1-1763 <JB>  
A:Cross-references: GB:M67507; NID:gl59648; PIDN:AAA18014.1; PID:gl59649  
C:Genetics:  
A:Introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfide  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:17-1763/Product: collagen alpha 2(IV) chain #status predicted <MAT>  
F:27-42/Domain: non-collagenous NHI #status predicted <NHI>  
F:43-1529/Domain: collagenous #status predicted <COL>  
F:157-199/Region: cell attachment (R-G-D) motif  
F:1830-1763/Domain: carboxyl-terminal nonhelical, NCI #status predicted <NC1>  
F:1530-1638/Domain: repeat NCI #status predicted <NC11>  
F:1639-1763/Domain: repeat NCI #status predicted <NC12>  
F:31,34,39,41,536,539/Disulfide bonds: interchain #status predicted  
F:126/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:1593-1599,1702-1709/Disulfide bonds: #status predicted  
Alignment Scores:  
Pred. No.: 2.5e-11 Length: 1763  
Score: 312.50 Matches: 315  
Percent Similarity: 31.65% Conservative: 60  
Best Local Similarity: 26.58% Mismatches: 430  
Query Match: 4.87% Indels: 381  
DB: 2 Gaps: 68  
us-09-759-143-110 (1-3410) x S16366 (1-1763)  
QY 2 GGAACAGCTGACGCGCTGGCTGCCGGTGACAGCGCGCTCGCCAGGATCTGAG 61  
Db 404 GlyProGlyLeuProGlyLeu---ProGlyLeuGluGlyLeuProGlyProLysGlyGlu 422  
QY 62 TGATGAGACGTGTCCCACTGAGTGTGCCCCACAGCAGAGTGTGTGAG-----CATGGG 115  
Db 423 LysGlyAspSerGly-----IleProGlyAlaProGlyValGlnGlyProGly 439  
QY 116 CTGAGAGCTGGACGGGACCAAGAGCTGCCAAGATGGCGCT-----GG 163  
Db 440 LeuAlaGlyProProGlyAlaLysGly-----GluProGly-ProArgGlyValAspG 457  
QY 164 CTGATTCTAGCAGATTGGCGGACGAGGAGAGGCCCGCTCTTGAGCAGAGCC 223  
Db 457 yGlnSerIleProGlyLeuProGlyLysAspGlyArgProGlyLeuAspGlyLeuProG 477  
QY 224 GAGACGAGCAGTGTCTGGAGTGCCT-----GAACGGCCCTCTGAG----- 263  
Db 477 yArgLysGlyGluMetGlyLeuProGlyValArgGlyProGlyAspSerLeuAsnG 497  
QY 264 ----CCCTACCGCGCTGCCACATGATGCTCAGAGGCTGTGGGTGACCGCGCTGCTGGCG 319  
Db 497 yLeuProGlyProProGlyPro-ArgGlyProGln-----GG 509

```

Db      829  roProGlyPheProGlyLeuLysGlyLysAspGlyIleProGlyAlaProGlyLeuProG 849
QY      1292 GTCTTCTCTCTGTCATGACCGGCTGCTGAGCGATTTCGCACTCGAGCAGCTATTG 1351
Db      849  lLeuGluGlyGlnArgGlyLeuProGlyValProGlyGlnLysGlyGlu-IleGlyLeu 868
QY      1352 GCCAGTGTGGCA-----GCCTTCTCTGCTGCTGCGGTGCCATGCTCTGCCAC 1402
Db      869  ProGlyLeuAlaGlyAlaProGlyPheProGlyAlaLysGlyLysGlyLeu-ProGly 888
QY      1403 AGTG-----TGCCCTGTGTGACAGCTTCAGCGCGCCTCACCGG 1441
Db      888  yLeuProGlyLysGluGlyProGlnGlyPro-----ProGlyGlnProGlyAlaProGly 906
QY      1442 TTCACCTTCTGAGCCCTCGAGCTTCGCTGCA-----1473
Db      906  yPheProGlyGln-LysGlyAspGlyLysGlyLeuProGlyLeuProGlyValSerGlyMetL 926
QY      1474 -----CACACTGGCTCTCTACCGCGGAGAGCAGGTGTCCTGCCAA-----A 1522
Db      926  yGlyAspThrGlyLeuProGlyValProGlyLeuAlaGlyAlaLysGlyGluProGlyP 946
QY      1523 TACCGAGGGGACTGGA-----GGTGTAGCAGTGAAGAC 1558
Db      946  hePro-GlyGlnLysGlyGlnProGlyPheProGlyValAlaGlyAlaLysGlyGlu--- 964
QY      1559 AGCCTGATGACACCTTCTGCTCCAGCGCTTACCGCGCTTCCCTCCCTAATGACAC 1618
Db      965  -----AlaGlyLeuProGly-----LeuProGlyAlaProGlyGlnLysGlyGlu 979
QY      1619 GTGGTGTGTGAGCAGTGGCTCTCTCCACCTCCAGCGCTCTCGCGGGCTCTGCC 1678
Db      980  GlnGlyLeuAlaGly-----LeuProGlyIleProGlyMetLysGlyAlaProGly 996
QY      1679 TGTGATGTCTCGTACGTGTGTGGTGGTGAAGCCACCGAGGC-----1722
Db      997  -----IleProGlyAlaProGlyGlnAspGlyLeuProGlyLeu 1009
QY      1723 -----CAGGTGTGTTCCGGCGGGGCGATCTGCTGCGCTGCG-----CATCTGGA 1770
Db      1010 ProGlyValLysGlyAspArgGlyPheAsnGlyLeuProGlyGlnLysGlyGluProGly 1029
QY      1771 TAGTGCCTT-----CTGTCTCTCCAGTGGCGCCATCCCTGTT 1809
Db      1030 ProAlaAlaArgAspGlyGluLysGlyGluProGlyLeuProGlyGlnPro-----1046
QY      1810 TATGGGTCTCATGTCAGCTCAGCGCTGCTGCTGCTATATGCTCTCTGCCGAGG 1869
Db      1047  ---GlyLeuArgGlyProGlnGlyProPro-----GlyLeuProGlyLeu 1060
QY      1870 CCTGGGTCTGTCGCCATTTACTTTGCTACACAGTAGTATTTGACAGCGCACTGGC 1929
Db      1061 ProGly-----1062
QY      1930 CAATACTCAGCTAGAAAACCTTCCAGCACATTGGGTGGAGGCGCTGCTCTACTGGGTC 1989
Db      1063  ---LeuLysGlyAspGlyGlnProGlyTyGlyAlaProGlyLeu-MetGlyGlu 1081
QY      1990 CCAGCTCCCGCTCTGTTAGCCCAATGCGGCTGCGCGCTGCGCGCAGTTCTGTTCG 2049
Db      1081 sGlyLeuPro-----GlyLeuProGlyLysPro-----1090
QY      2050 TGCCAAAGTAATGTGGCTCTCTGCTGCCACCTGTGCTGCTGAGGTGCTAGCTGCCACAG 2109
Db      1090 -----1090
QY      2110 CTGGGGCTGGGCGTCC-----TCCTCTCTCTCCCAAGTCTCTA---GGGTGCTGCTGACT 2163
Db      1091  -----GlyArgProGlyAlaPro-GlyProLysGlyLeuAspGlyAlaProGlyP 1107
QY      2164 GGAGGCTTCCAAAGGGGTTTCAGTCTGACATTATACAGGGAGGCGCAAGAGGCTCCATG 2223

```

```

Db      1107 heProGlyLeuLysGlyGluAla-----GlyLeuPro---GlyAlaProG 1121
QY      2224 CACTGGAATCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTACACGATAGCCTC 2283
Db      1121 lLeuProGlyGlnAspGlyLeuProGlyLeuProGlyGlnLysGlyGluSerGlyPhe 1141
QY      2284 CTAGTTTGAGACACACCTAGAGAAGGTTTGGGAGCTGTAATAACTCAGTCACCTGGTT 2343
Db      1141 roGly-----GlnProGlyLeuValGly-----ProProGlyL 1152
QY      2344 TCCCATCTCTAAGCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTCATGGGAGTTT 2403
Db      1152 euProGlyLys-----MetGlyAlaP 1159
QY      2404 CTAGGATGAACACTCTCCATGGGATTTCAACATATGACTTATTTGTAGGGAAG-----2459
Db      1159 roGlyIleArg-----GlyGluLysGlyAspAlaGlyLeuProGlyLeuProGlyGluArg 1178
QY      2460 -----AGTCTGAGGGGCAACACACAGAACAGGTCCTCCCTCAGCCACAGCACTCT 2511
Db      1178 lLeuAspGlyLeuProGlyGlnLysGlyGluAlaGlyPhePro-----1192
QY      2512 CTTTTTGTGATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTGTGCTCTCTG 2571
Db      1193 -----GlyAlaProGlyLeuPro-----GlyProValGlyProLysG 1205
QY      2572 TTGCCATCAGACAGACAGGCACTTTAAATATTTAACTTATTTATTTAACTAAGTAGAG 2631
Db      1205 lSerAlaGlyAlaProGly-----PheProGlyLeuLysG 1217
QY      2632 GGAATCCATGCTAGCTTTCTGTTGTTGTTCTAATATTTGGTAGGTGGGGGATCC 2691
Db      1217 lGluProGlyLeuProGlyLeu-----GluGlyGlnP 1228
QY      2692 CAACAATCAGGTCC-----2705
Db      1228 roGlyProArgGlyMetLysGlyGluAlaGlyLeuProGlyAlaProGlyArgAspGlyL 1248
QY      2706 ---CCTGAGATAGCTGATTTGGGTGATTCGTCAGAACTCTCTCTCTCTCTCTCTG 2763
Db      1248 euProGlyLeuProGlyMetLysGlyGlu-----AlaGlyLeuProGlyLeuProGly 1266
QY      2764 GCCCCCCAAATGCTTAACCCAGGACCTTGAATTTCTACTATCCCAATGATAATTC 2823
Db      1266 lProGlyLysSerIleThrGlyProLysGlyAsnAlaGly-LeuProGlyLeuProGly 1285
QY      2824 AAATGCTGTACCCAGGTAGGTGTTGAAGGAAGGTAGAGGT-----2868
Db      1286 LysAspGlyLeuProGlyLeuProGlyLeuProGlyLysGlyGluProGlyLysProGlyTyrAla 1305
QY      2869 -----GGGCT 2874
Db      1306 GlyAlaAlaGlyIleLysGlyGluProGlyLeuProGlyIleProGlyAlaLysGlyGlu 1325
QY      2875 TCAGGTCTCAAGGCTTCCCTAACCC-----CCCTCTCTCTCTCTCTCTCTCTG 2916
Db      1326 ProGlyLeuSerGlyIleProGlyLysArgGlyAsnAspGlyIleProGlyLysProGly 1345
QY      2917 CCAGCTGTGTTCCCGCCCTCTCTCA-----CTCCCTCTCTCTCTCTCTCTCTCTC 2955
Db      1346 ProAlaGlyLeuProGlyLeuProGlyMetLysGlyGluSerGlyLeuProGlyProGln 1365
QY      2956 TCTAGGACTGGGTGATGAAGGCACTGCCAAAATTTCCCTACCCCACTTCC 3012
Db      1366 GlyProAlaGlyLeu---ProGlyLeuProGlyLeuLysGlyGluProGlyLeuPro 1383

```

## RESULT 35

A54849  
collagen alpha 1(VII) chain precursor - human  
N:Alcinate names: procollagen alpha 1(VII) chain  
C:Species: Homo sapiens (man)  
C:Date: 04-Nov-1994 #sequence\_revision 04-Nov-1994 #text\_change 20-Sep-1999  
C:Accession: A54849; PH0844; S16316; I56328; A30296; I84686



R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.  
J. Biol. Chem. 269, 20256-20262, 1994  
A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(VII)  
A:Reference number: A54849; MUID:94327588; PMID:8051117  
A:Accession: A54849  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-2944 <CHR>  
A:Cross-references: GB:102870; NID:987124; PIDN:AAA75438.1; PID:g987125  
R:Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.  
Biochem. Biophys. Res. Commun. 183, 958-963, 1992  
A:Title: Molecular cloning and characterization of type VII collagen cDNA.  
A:Reference number: PH0844; MUID:92231902; PMID:1567409  
A:Accession: PH0844  
A:Molecule type: mRNA  
A:Residues: 'EFR', 340-475, 'RALSTASHSTLWRAWRWPCNRGSHWTRAAACEPCNRPASHAARAG', 524-528, 'C',  
A:Cross-references: DDBJ:D11152; DDBJ:D13694; NID:9453698; PIDN:BAA02853.1; PID:g453699  
A:Experimental source: keratinocyte  
A:Note: the authors translated the codon ACC for residues 394 and 397 as Tyr  
R:Parente, M.G.; Chung, L.C.; Ryyanen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mat  
Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991  
A:Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.  
A:Reference number: S16316; MUID:91334380; PMID:1871109  
A:Accession: S16316  
A:Molecule type: mRNA  
A:Residues: 815-892, 'E', 894-1439 <PAR>  
A:Cross-references: GB:M65158; GB:S49017; NID:g180914; PIDN:AAA96439.1; PID:g180915  
A:Experimental source: keratinocyte  
R:Gammon, W.R.; Abernethy, M.L.; Padilla, K.M.; Prisyanyh, P.S.; Cook, M.E.; Wright, J.;  
J. Invest. Dermatol. 99, 691-696, 1992  
A:Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot  
A:Reference number: I56328; MUID:93107742; PMID:1469284  
A:Accession: I56328  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 'EFR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>  
A:Cross-references: GB:S51236; NID:g262308; PIDN:AAB24637.1; PID:g262309  
R:Seitzer, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.  
J. Biol. Chem. 264, 3822-3826, 1989  
A:Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagena  
A:Reference number: A30296; MUID:89139437; PMID:2537292  
A:Accession: A30296  
A:Molecule type: protein  
A:Residues: 'A', 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'E', 2032, 'C', 2034-2041;  
A:Note: two reported peptides cannot be reliably located  
R:Greenspan, D.S.  
Hum. Mol. Genet. 2, 273-278, 1993  
A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous  
A:Reference number: I48103; MUID:93271985; PMID:8499916  
A:Accession: I48103  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 2395-2871, 'S', 2873-2944 <RE2>  
A:Cross-references: GB:L06862; NID:g388713; PIDN:AAA89196.1; PID:g388714  
R:Christiano, A.M.; Ryyanen, M.; Uitto, J.  
Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994  
A:Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs  
A:Reference number: A55255; MUID:94224777; PMID:8170945  
A:Contents: annotation  
A:Comment: prolines and lysines at the third position of the tripeptide repeating unit  
ed and subsequently O-glycosylated.  
C:Genetics:  
A:Gene: GDB:COL7A1; EBR1; EBD1; EB  
A:Cross-references: GDB:128750; OMIM:120120  
A:Map position: 3p21.3-3p21.3  
A:Note: defects in this gene can result in dominant and recessive dystrophic epidermolys  
C:Complex: there are 118 introns  
A:Note: type VII collagen is probably a homotrimer  
C:Function:  
A:Description: structural component of extracellular polymer associated with anchoring  
C:Superfamily: unassigned collagens; animal Kunitz-type proteinase inhibitor homology;  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli  
F:1-16/Domain: signal sequence #status predicted <STG>

F:17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>  
F:17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>  
F:36-201/Domain: von Willebrand factor type A repeat homology <VWA1>  
F:231-318/Domain: fibronectin type III repeat homology <FN1>  
F:327-413/Domain: fibronectin type III repeat homology <FN2>  
F:414-502/Domain: fibronectin type III repeat homology <FN3>  
F:508-593/Domain: fibronectin type III repeat homology <FN4>  
F:598-683/Domain: fibronectin type III repeat homology <FN5>  
F:686-771/Domain: fibronectin type III repeat homology <FN6>  
F:776-862/Domain: fibronectin type III repeat homology <FN7>  
F:864-952/Domain: fibronectin type III repeat homology <FN8>  
F:954-1045/Domain: fibronectin type III repeat homology <FN9>  
F:1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>  
F:1170-1172/Region: cell attachment (R-G-D) motif  
F:1189-1253/Region: cysteine/proline-rich  
F:1254-2783/Region: interrupted helical  
F:1334-1336/Region: cell attachment (R-G-D) motif  
F:2553-2555/Region: cell attachment (R-G-D) motif  
F:2876-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>  
F:2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
F:337-786, 1109/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:2167, 2176, 2185, 2188, 2664, 2667, 2673/Modified site: 4-hydroxyproline (Pro) #status ex  
F:2625, 2631/Modified site: 5-hydroxylysine (Lys) #status experimental  
F:2625, 2631/Binding site: carboxylate (Lys) (covalent) #status experimental  
F:2634, 2802, 2804/Disulfide bonds: interchain #status predicted

Alignment Scores:  
Pred. No.: 2,5e-11 Length: 2944  
Score: 312.50 Matches: 260  
Percent Similarity: 31.26% Conservative: 37  
Best Local Similarity: 27.37% Mismatches: 362  
Query Match: 5.02% Indels: 291  
DB: 2 Gaps: 53

US-09-759-143-110 (1-3410) x A54849 (1-2944)

QY 2489 CCTGTTCTTGTGTGTGTTG-----CCCTCAGGACTCTTCCCTACAAATAAGT 2442  
DB 1158 ProGlyValMetValLeuValAspGluProLeuArg----- 1170  
QY 2441 CATATGTTCAATCCATCCATGGAGGAGTGTTTCATCTAGTAACTCCATGCAGAGACTACA 2382  
DB 1171 -----GlyAspIlePheSerPro----- 1176  
QY 2381 TTAACAGGAGCTGCAGGTTTAAGGGG----- 2357  
DB 1177 IleArgGluAla-GlnAlaSerGlyLeuAsnValValMetLeuGlyMetAlaGlyAlaAs 1196  
QY 2356 -----CTTAGAGATGGGAACACAGGT---GACTGAGTT-----TATTACGC 2319  
DB 1196 pProGluGlnLeuArgArgLeuAlaProGlyMetAspSerValGlnThrPhePheAlaVa 1216  
QY 2318 TCCCAAAAACCCCTTCTCTA-----GGTGTGTCTCACTAGGAGGCTAGCTGTTAACCCCT 2265  
DB 1216 lAspAspGlyProSerLeuAspGlnAlaValSerGlyLeuAlaThrAlaLeuCysGlnAl 1236  
QY 2264 GAGCCCTGGCTAATCCACCTGCAGAGTCCCGCCATTCACGTCATGGAGCCCTCTCGGCT 2205  
DB 1236 aserPheThrGlnProArgProGluPro---CysProValTyrCysProLysGly--- 1254  
QY 2204 CCTGTATTAAGTCCAGACTGAACCCCTTGGAGGCCCTCCAGCGAGCCCTAGAGAG 2145  
DB 1255 -----GlnLysGlyGluProGlyG1 1261  
QY 2144 CTGGGAGAGAGAGA-----GGGAGCCGCCAGCCCGAGCTGTCACCTAGGACCTC 2091  
DB 1261 uMetGlyLeuArgGlyGlnValGlyProGlyAspProGlyArgThrG1 1281  
QY 2090 AGCAGACACGGTGGCGAGAGAGACACATTTGTCAGCAGCAACAGAAATGCGGCC 2031  
DB 1281 yAlaPro-----GlyProGlnGlyProPro-----GlySer-AlaThrAlaLysGlyG 1297

QY	2030	AGCCCCGGACGCCCATCGGGCTAACAGGACGGG---GGAGCTGGGACCCCACTGAGCGCAGG	1979
Db	1297	luArgGlyPhePro--GlyAlaAspGly--ArgProGlySerProGlyArgAlaGlyAsn	1315
QY	1973	CCCTCCACCCCAATGTGCTGGAAGTTTCTACGCTGAGTATTTGGCCANGTCCTCTCTGT	1914
Db	1316	ProGlyThrProGlyAla-	1321
QY	1913	CAATACTACTCTGTGTAGCAAAGTAATGCGCACCAGACCCAGGCCTGGCGGACACACCA	1854
Db	1322	-----ProGlyLeuLysGlySerPro	1328
QY	1953	TATAGGCAGTGCACAGACTGGCTGAGCTGGACAAATGAGCCCATAAACAGGATGGGCCCA	1794
Db	1329	GlyLeuProGlyProArgGlyAspProGlyGluArgGlyProArgGlyProLysGlyGlu	1348
QY	1793	CCTGGGACAGCAGGAAGG-----CACTATCCAGGATGGCAGGTCCAGGCAGA	1746
Db	1349	ProGlyAlaProGlyGlnValIleGlyGlyGluGlyProGlyLeuProGlyArgLysGly	1368
QY	1745	TGCCCGGCCCGGAA---CCACCTCGCTCGGTGGGTCAACCACACACACGCTACGC	1689
Db	1369	AspProGlyProSerGlyProProGlyProArg-Gly-----ProLeuG1	1383
QY	1688	AGACATCACAGGCAGAGCCCC-----GCAGGCGCGG	1656
Db	1383	YAspProGlyProArgGlyProProGlyLeuProGlyThrAlaMetLysGlyAspLysG1	1403
QY	1655	GTGAGGTGGGACGACGCCACTGCCCTCCAGCACCCACGCTGTCCATTGAGGAAGGAGCTC	1596
Db	1403	YAspArg-GlyGluArgGly---ProProGlyPro-----GlyGluG	1416
QY	1595	CAGGCTTAGGGCTGCGCAGGAAGCTGCTCATCAGGCTGTCTCTACTGTAGCACCTCCAG	1536
Db	1416	LycGlyIleAlaProGly-----	1421
QY	1535	TGTCCTCCCTGGTATTTGGCAGGAACACCTGCTTCTCCGGTGTGTAGAGGGAGGCAGTG	1476
Db	1422	--GluPro-GlyLeu---ProGlyLeuProGlySerProGlyProGlnGly---ProVal	1438
QY	1475	TGTAGGCGCATCTGCAGGGCTGAGAAGGTGAACCCGGTGGAGGGCGGTGAAGCTGTCA	1416
Db	1439	GlyProGlyLysLysGlyGluLysGlyAspSerGluAspGlyAlaProGlyLeuPro	1458
QY	1415	CCAGCGCCACACTGTGGGACAGGCATGTGGCACCCGCGCACGCCACAGGAAAGCTG----	1362
Db	1459	GlyGlnProGlySerProGlyGluGlnGlyProArgGlyProProGlyAlaIleGlyPro	1478
QY	1361	-----CCACACTGGCCAAATAGACTGCTCGAGTCCGAGTCCGGAATCGCTGCCACCCCGGT	1311
Db	1479	LysGlyAspArgGlyPheProGlyProLeuGlyGluAlaGlyGluLysGlyGluArgGly	1498
QY	1310	CCATGACCTAGAGAGACCGAGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCA	1251
Db	1499	Pro-----ProGly-----ProAlaGlySerArg---GlyLeuPro	1509
QY	1250	TCCGAAGCCTTCATCATAGTGTCTCCGGGCTCGGTGCCGGCTCAGCTGTGGGCACGC	1191
Db	1510	GlyValAlaGlyArgProGlyAlaLysGlyProGluGlyProProGlyProThrGlyArg	1529
QY	1190	CCTGTTACAGCCCTCGCCCAAGAAATCCGTGTAAACAGCCTGAAGTTCATGATGCCA	1131
Db	1530	GlnGlyGluLysGlyGluProGlyArgProGlyAspProAla-----ValVal	1545
QY	1130	TCCAGCTGCACAGCTAGCCACAGAGCCGGCGCAGGTCGGGGCATGCGGC-----	1077
Db	1546	GlyProAlaValAlaGlyProLysGlyGluLysGlyAspValGlyProAlaGlyProArg	1565
QY	1076	AGCACAGCTGTTCAGCCGGGGAACAGGGGCCCGCCAGGTTCCGGAAACCCAGCGGCC	1017
Db	1566	GlyAlaThrGlyValGlnGlyArgGlyProProGlyLeuValLeuProGlyAspPro	1585
QY	1016	GCATGGACAGCAGTGGGCGCACAAAGGAGGGGGCCGACAGGCCCTTCTGCTGGCTCGTGG	957

Db 1894 roValGlyProGlyGlnGlyPheProGlyValProGlyGlyThrGlyProLysGlyA 1914

QY 56 ATCTCTGGC-----CGAGCGCGCGGC 36  
||| |||  
1914 spArgGlyGluThrGlySerLysGlyGluGlnGlyLeuProGlyGluArgGlyLeuArgG 1934  
||| |||  
QY 35 TGTCAACCGCGAGCC 22  
||| |||  
Db 1934 lyGluProGlySer 1938

RESULT 36  
CGHU2V  
collagen alpha 2(V) chain precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Jul-1989 #sequence\_revision 28-Jul-1995 #text\_change 31-Dec-2000  
C:Accession: A31427; A54555; S43643; A25874; I55239; I59025; A25374; A30017  
R:Woodbury, D.; Benson-Chanda, V.; Ramirez, F.  
J. Biol. Chem. 264, 2735-2738, 1989  
A:Title: Amino-terminal propeptide of human pro-alpha2(V) collagen conforms to the struc  
A:Reference number: A31427; MUID:89123368; PMID:2914927  
A:Accession: A31427  
A:Molecule type: mRNA  
A:Residues: 1-463 <WOO>  
A:Cross-references: GB:J04478; NID:gl79697; PIDN:AAA51859.1; PID:gl79698  
A:Experimental source: placenta  
R:Greenspan, D.S.; Lee, S.T.; Lee, B.S.; Hoffman, G.G.  
Gene Expr. 1, 29-39, 1991  
A:Title: Homology between alpha2(V) and alpha1(III) collagen promoters and evidence for  
A:Reference number: A54555; MUID:92314691; PMID:1820205  
A:Accession: A54555  
A:Molecule type: DNA  
A:Residues: 1-32 <GRE>  
A:Cross-references: GB:M58529; NID:gl80834; PIDN:AAC41699.1; PID:G553235  
R:Moradi-Ameli, M.; Rousseau, J.C.; Kleman, J.P.; Champilaud, M.F.; Boutillon, M.M.; Ber  
Eur. J. Biochem. 221, 987-995, 1994  
A:Title: Diversity in the processing events at the N-terminus of type-V collagen.  
A:Reference number: S43642; MUID:94237164; PMID:8181482  
A:Accession: S43643  
A:Molecule type: protein  
A:Residues: 288-291, 'P', 293-294, 'X', 296-297, 606, 'X', 608-617 <MOR>  
R:Well, D.; Bernard, M.; Gargano, S.; Ramirez, F.  
Nucleic Acids Res. 15, 181-198, 1987  
A:Title: The pro alpha 2(V) collagen gene is evolutionarily related to the major fibrill  
A:Reference number: A25874; MUID:87146331; PMID:3029669  
A:Accession: A25874  
A:Molecule type: mRNA: DNA  
A:Residues: 388-1496 <WEI>  
A:Cross-references: GB:X04758; NID:g29588; PIDN:CAA28454.1; PID:gl340175  
A:Experimental source: rhabdomyosarcoma cell line  
R:Myers, J.C.; Loidl, H.R.; Stollé, C.A.; Seyer, J.M.  
J. Biol. Chem. 260, 5533-5541, 1985  
A:Title: Partial covalent structure of the human alpha 2 type V collagen chain.  
A:Reference number: I55239; MUID:85182703; PMID:2985598  
A:Accession: I55239  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1002-1226 <RE2>  
A:Cross-references: GB:M10956; NID:gl80427; PIDN:AAA52007.1; PID:gl80428  
R:Note: part of this sequence were determined by protein sequencing  
R:Emanuel, B.S.; Cannizzaro, L.A.; Seyer, J.M.; Myers, J.C.  
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985  
A:Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm  
A:Reference number: I59025; MUID:85216505; PMID:3858826  
A:Accession: I59025  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1003-1034 <RES>  
A:Cross-references: GB:M11135; NID:gl79693; PIDN:AAA51857.1; PID:gl79694  
R:Note: part of this sequence were determined by protein sequencing  
R:Myers, J.C.; Loidl, H.R.; Seyer, J.M.; Dion, A.S.  
J. Biol. Chem. 260, 11216-11222, 1985  
A:Title: Complete primary structure of the human alpha-2 type V procollagen COOH-termina  
A:Reference number: A25374; MUID:85289337; PMID:2411731

A:Accession: A25374  
A:Molecule type: mRNA  
A:Residues: 1227-1417, 'T', 1419-1437, 'S', 1439-1496 <WYE>  
A:Cross-references: GB:M11718; NID:gl80912; PIDN:AAA52058.1; PID:gl80913  
R:Experimental source: normal fibroblasts  
R:Tsipouras, P.; Schwartz, R.C.; Liddell, A.C.; Salkeld, C.S.; Weil, D.; Ramirez, F.  
Genomics 3, 275-277, 1988  
A:Title: Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2, located  
A:Reference number: A30017; MUID:89138450; PMID:3224983  
A:Accession: A30017  
A:Molecule type: DNA  
A:Residues: 1449-1463, 'E', 1465-1495, 'A' <TSI>  
A:Cross-references: GB:J03051; NID:gl79695; PIDN:AAA51858.1; PID:gl79696  
A:Note: the authors translated the codon GAA for residue 1460 as Gln, and GAG for res  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating uni  
are 5-hydroxylated and subsequently O-glycosylated.  
C:Comment: The amino-terminal propeptide domain appears not to be completely cleaved.  
C:Genetics:  
A:Gene: GDB:COL5A2  
A:Cross-references: GDB:I19064; OMIM:120190  
A:Map position: 2q31-2q31  
A:Introns: 33/1; 812/3; 830/3; 848/3; 884/3; 902/3; 922/3; 974/3; 1064/3; 1064/3; 144  
C:Complex: type V collagen may be a homotrimer of alpha 1(V) chains (see PIR:CGHULV).  
alpha 2(V) chain and one alpha 3(V) chain, initially linked by disulfide bonds among  
length, is formed with desmosine cross-links made from lysine and allysine residues  
C:Function:  
A:Description: structural component of extracellular fibrous polymer associated with  
A:Note: may play a role in controlling the lateral growth of collagen I fibrils  
C:Superfamily: collagen alpha 1(i) chain; fibrillar collagen carboxyl-terminal homolo  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxypr  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-1250/Product: collagen alpha 2(V) chain #status predicted <MAT>  
F:27-193/Domain: amino-terminal propeptide (uncleaved) #status predicted <NPP>  
F:27-108/Region: nonhelical  
F:40-99/Domain: von Willebrand factor type C repeat homology <WVC>  
F:109-186/Region: helical  
F:187-208/Region: nonhelical  
F:209-1225/Region: helical  
F:503-505/Region: cell attachment (R-G-D) motif  
F:941-943/Region: cell attachment (R-G-D) motif  
F:1064-1066/Region: cell attachment (R-G-D) motif  
F:1067-1069/Region: cell attachment (R-G-D) motif  
F:1097-1099/Region: cell attachment (R-G-D) motif  
F:1124-1126/Region: cell attachment (R-G-D) motif  
F:1133-1135/Region: cell attachment (R-G-D) motif  
F:1225-1250/Region: carboxyl-terminal nonhelical  
F:1251-1496/Domain: fibrillar collagen propeptide #status predicted <CPP>  
F:267/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic  
F:193-194/Cleavage site: Ala-Gln (procollagen N-endopeptidase) #status predicted  
F:194/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predi  
F:201/Modified site: allysine (Lys) #status predicted  
F:290-293, 296, 608, 614, 1004, 1007, 1013, 1028, 1034/Modified site: 4-hydroxyproline (Pro)  
F:299, 1139/Modified site: 5-hydroxylysine (Lys) #status predicted  
F:299, 1139/Binding site: carboxylate (Lys) (covalent) #status predicted  
F:1025/Modified site: 5-hydroxylysine (Lys) #status experimental  
F:1259, 1397/Cleavage site: Glu-Asp (procollagen C-endopeptidase) #status predicted  
F:1259, 1397/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:1293, 1399, 1325/Disulfide bonds: interchain #status predicted  
F:1333-1494, 1402-1447/Disulfide bonds: #status predicted

Alignment Scores:  
Pred. No.: 2, 68e-11 Length: 1496  
Score: 312.00 Matches: 226  
Percent Similarity: 30.80% Conservative: 37  
Best Local Similarity: 26.46% Mismatches: 255  
Query Match: 4.86% Indels: 336  
DB: 1 Gaps: 51

US-09-759-143-110 (1-3410) x CGHU2V (1-1496)

QY 26 CCGGGTGCAC-----AGCCGCGCGCTTCGG-----CCAGGA 55  
||||||| ||| |||||

Db 242 ProGlyAspProGlyProMetGlyProIleGlySerArgGlyProGluGlyProProGly 261  
 QY 56 TCTAGTATGATGAGAGCTGTCCTCCACTGAGG----- 85  
 Db 262 LysProGlyGluAspGlyGluProGlyArgAsnGlyAsnProGlyGluValGlyPheAla 281  
 QY 86 -----TGCCCCACAGCAGCAGGCTGTGAGCATGGCTGAGAAC 124  
 Db 282 GlySerProGlyAlaArgGlyPheProGlyAlaProGlyLeuPro---GlyLeuLysGly 300  
 QY 125 TGGACCGGCACCAAGGCGTG----- 145  
 Db 301 HisArgGlyHisLysGlyGluGlyProLysGlyGluValGlyAlaProGlySerLys 320  
 QY 146 -----GCAGAAATGGCGCCTGG-----CTGATTCCTAGCAGTTG 181  
 Db 321 GlyGluAlaGlyProThrGlyProMetGlyAlaMetGlyProLeuGlyProArgGlyMet 340  
 QY 182 GCGGAGCAAGAGGAGAGCGCGAGCTCTGGAGCAGAGCGAGCAAGCACTTCTGG 241  
 Db 341 ProGlyGluArg-GlyArgLeuGlyProGlnGlyAlaProGlyGlnArgGlyAlaHisGly 360  
 QY 242 AGTGCTGAACGCGCC----- 257  
 Db 360 yMetProGlyLysProGlyProMetGlyProLeuGlyIleProGlySerSerGlyPhePr 380  
 QY 257 ----- 257  
 Db 380 oGlyAsnProGlyMetLysGlyGluAlaGlyProThrGlyAlaArgGlyProGluGlyPr 400  
 QY 258 -----CCTGAGCCTACCGCCTGGCCCTACTATGTT-----CCAG 292  
 Db 400 oGlnGlyGlnArgGlyGluThrGlyProProGlyProVal-GlySerProGlyLeuProG 420  
 QY 293 AGGTGTGGTGAGCGCGCTGCTGGCAGCCGAA-----AGCCAGCTCTTG 340  
 Db 420 yAlaIleGlyThrAspGlyThrProGlyProLysGlyProThrGlySerProGlyThrS 440  
 QY 341 CTGTGCAACCTGTAACTTTGGCTGGAGGTGTGTTGGCCGAGCATCACTATGTG 400  
 Db 440 erGlyProProGlySerAlaGlyProProGlySerProGlyProGln----- 455  
 QY 401 CGCGCTCTGCTGGAAGTGGGGTAGAGGAGAGTTCATGACCATGGTCTGGGCATT 460  
 Db 456 --GlySerThrGlyProGlnGlyAsnSerGlyLeuProGlyAsp-----ProGlyPheL 473  
 QY 461 GGTCCAGTGTGGCGCTGTGTCCCGCTCTAGGCTCAGCAGTGCACCTAGTGGC- 519  
 Db 473 ysGlyGluAlaGlyProLysGlyGluProGlyProHisGlyIleGlnGlyProIleGlyP 493  
 QY 520 -----TGACGCTATGGCCGCG- 537  
 Db 493 roProGlyGluGlyLysArgGlyProArgGlyAspProGlyThrLeuGlyProProG 513  
 QY 538 --CGGCGCTTCATCTGGGCATCTC-----CTTGGGCATCTCT---GCTGAGCCTCTTT 586  
 Db 513 lyProValGlyGluArgGlyAlaProGlyAsnArgGlyPheProGlySerAspGlyLeuP 533  
 QY 587 CTATCCCAAGGCGCGCTGCTAGCAGGCTGCTGCCCGGATCCAGCCCGCTGGAG 646  
 Db 533 roGlyProLysGlyAlaGlnGlyGluArgGlyProVal---GlySerSerGlyProLysG 552  
 QY 647 CTGGCACTGCTCATCTTGGGCGT---GGGCGTGTGAGCTCTGTGCGCAGGTGTGCTTC 703  
 Db 552 lySerGlnGlyAspProGlyArgProGlyGluProGlyLeu-----ProGlyAlaArgG 570  
 QY 704 ACTCCACTGGAGGCGCTGCT-----CTCTGACCTCTTCCGGGAC 742  
 Db 570 lyLeuThrGlyAsnProGlyValGlnGlyProGluGlyLysLeuGlyProLeuGlyAlaP 590  
 QY 743 CCGGA-----CCACTGTGCGCCAGCGCTACTCTGTCTATGCTTTCATGATCAGT 790  
 Db 590 roGlyGluAspGlyArgProGlyProGly-----S 601

QY 791 CTTGGGGGCTGCTGGCTACCTCCTCCTGCTCCATTGACTGGACACCACTGCTGCTGGCC 850  
 Db 601 erIleGlyIleLysGlyGlnProGlyThr-----MetGly-----LeuProGlyP 616  
 QY 851 CCCTA-----CCTGGCACCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895  
 Db 616 roLysGlySerAsnGlyAspProGlyLysProGlyGluAla-----GlyAsnP 632  
 QY 896 CTATCTTCTCCTACCTGCTAGCAGCAGCAGCTGCTGCTGCTGAGGAGGAGGCTGGCC 955  
 Db 632 roGlyValProGlyGlnArgGly-----AlaProGlyLysAspGlyLysValGlyP 649  
 QY 956 CCAC-----CGACCCAGCAGAGGGCTGTGCGCCCTCC 991  
 Db 649 rofGlyProProGlyProProGlyLeuArgGlyGluArgGlyGluGlnGlyPro----- 567  
 QY 992 TTGTCGCCCCACCTGCTGCTCATCCGCGCGCTGCTTTCGGAACCTGGGCGCCCTG 1051  
 Db 668 -----ProGlyProThrGlyPhe-----GlnGlyHisProG 678  
 QY 1052 CTTCCCGGCTGACACGAGCTGCTGCTGCGCATGCCCGCACCTGCGCGGCTCTTCGTG 1111  
 Db 678 lyProProGlyProProGly----- 684  
 QY 1112 GCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTACACGGATTTCTG 1171  
 Db 685 -----GluGly-----GlyLysProG 690  
 QY 1172 GCGAGGGGCTGTACACAGGCGTGCAGAGCTGAGCGGCGCACCGAGCGCGGAGACAC 1231  
 Db 690 lyAspGlnGlyValProGlyGlyProGlyAla---ValGlyProLeuGlyPro----- 706  
 QY 1232 TATGATGAGCGCTTCGGATGGGACGCTGG-----GCTGTTCCTGCTGAGTGGCGCATC 1285  
 Db 707 -----ArgGlyGluArgGlyAsnProGlyGluArgGlyGluProGlyIleThrGlyL 724  
 QY 1286 TCCT-----GGTCTTCTCTGTCATGACCGCTGCTGTCAGCGATTCGGCACT 1336  
 Db 724 euProGlyGluLysGlyMetAlaGlyHisGlyProAspGly----- 738  
 QY 1337 CGAGCAGCTATTTGGCCACTGTGGCAGCTTTCCTGCTGCTGCGCGTCCACATGCTCTG 1396  
 Db 739 -----ProLysGly-----SerProG 744  
 QY 1397 TCCACAGCTGTGGCGCTGTGACAGCTTCAGCGCGCTCAGCGGTTTACCTTCTCAGCC 1456  
 Db 744 lyProSerGlyThrProGlyAspThr----- 752  
 QY 1457 CTGAGATCTGCTTACACTGCTCCTTCTACACCGGAGAGCAGGTGTTCTCTG 1516  
 Db 753 --GlyProProGlyLeuGln---GlyMetProGlyGluArgGlyIleAlaGlyThrProG 771  
 QY 1517 CCCAATACCGAGGCGACACTGGAGGTGTAGCAGTGAAGCAGCAGCTGTATGACACG--- 1573  
 Db 771 ly-Pro---LysGlyAspArgGlyIleGlyGluLysGlyAlaGlnGlyThrAlaGly 789  
 QY 1574 -----TTCTGCGCAGCGCTAAGCCTGGAGCTCCCTTCCCTTAAATGGACAC 1618  
 Db 790 AsnAspGlyAlaGlyGlyLeuProGlyProLeu---GlyProProGlyProAlaGlyLeu 808  
 QY 1619 GTGGGTGCTGGAGGAGT-----GGCTGCTCCCACTCCA----- 1654  
 Db 809 LeuGlyGluLysGlyGluProGlyProArgGlyLeuValGlyProProGlySerArgGly 828  
 QY 1655 ---CCGCGCTCTGCGGCGCTGCTGCTGTGATGCTCGCTACGTGTG----- 1699  
 Db 829 AsnProGlySerArgGlyGluAsnGlyProThrGlyAlaValGlyPheAlaGlyProGln 848  
 QY 1700 -----GTGTGGGTGAGCGCCACCGAGCGGCTGTTCCGGGC 1738  
 Db 849 GlySerAspGlyGlnProGlyValLysGlyGluProGlyGlu-Pro-GlyGlnLysGlyA 868

QY 1739 CGGGGCATCTGCTGACCTCGC-----CATCTGGATAGTCTCTCTGCTG 1786  
 Db |||||  
 868 spAlaGlySerProGlyProGlnGlyLeuAlaGlySerProGly----- 882  
 QY 1787 TCCAGGTGGCCCATCTCTGTTATGGCTCCAT-----TGTCCAGTCTAGC 1834  
 Db |||||  
 883 --ProHisGlyProAsnGlyValProGlyLeuLysGlyGlyArgGlyThrGlnGlyPro 902  
 QY 1835 CAGTCTGCTACTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1894  
 Db |||||  
 902 ro-----GlyAlaThrGlyPheProGlySerAla----- 911  
 QY 1895 GCTACACAGTAGTATTGACAAAGAGCGACTTGGCCAAATACTACGCGTAGAAACTTCC 1954  
 Db |||||  
 911 ----- 911  
 QY 1955 AGCACATGTTGGTGGAGGCGCTGCTCTACTGGGTCCAGCTCCCGCTCTCTTACGCC 2014  
 Db |||||  
 912 -----GlyArgValGlyProProGlyProAlaGlyAla-ProGlyProAlaGlyPro 928  
 QY 2015 ATGGGCTGCGGGC-----TGGCCGCCA 2038  
 Db |||||  
 929 LeuGlyGluProGlyLysGluGlyProPro 938

## RESULT 37

149607

Procollagen type V alpha 2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 13-Aug-1999

C:Accession: I49607

R:Andrikopoulos, K.; Suzuki, H.R.; Solursh, M.; Ramirez, F.

Dev. Dyn. 195, 113-120, 1992

A:Title: Localization of pro-alpha 2(V) collagen transcripts in the tissues of the devel

A:Reference number: I49607; PMID:93214071; PMID:1297453

A:Accession: I49607

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-1497 &lt;RES&gt;

A:Cross-references: GB:L02918; NID:g309180; PIDN:AAA37440.1; PID:g309181

C:Genetics:

A:Gene: Col5a-2

C:Superfamily: collagen alpha 1(1) chain; fibrillar collagen carboxyl-terminal homology;

F:39-98/Domain: von Willebrand factor type C repeat homology &lt;WMC&gt;

F:1270-1497/Domain: fibrillar collagen carboxyl-terminal homology &lt;FCC&gt;

## Alignment Scores:

Pred. NO.:	2.88e-11	Length:	1497
Score:	311.50	Matches:	225
Percent Similarity:	30.02%	Conservative:	32
Best Local Similarity:	26.29%	Mismatches:	258
Query Match:	4.85%	Indels:	341
DB:	2	Gaps:	47

US-09-759-143-110 (1-3410) x I49607 (1-1497)

QY 20 CTGGCTCGGGTGACAGCCGCGCTCGG-----CCAGGATCTGATGATGAGAGCTG 73  
 Db |||||  
 246 MetGlyProLeuGlySerArgGlyProGluGlyProGlyLysProGlyGluAspGly 265  
 QY 74 TCCCACTAGG----- 85  
 Db |||||  
 266 GluProGlyArgAsnGlyAsnThrGlyGluValGlyPheSerGlySerProGlyAlaArg 285  
 QY 86 ---TCCGCCACAGCAGGAGTGTGACATGGCTGAGAGCTGGACCGCCACCAAGGG 142  
 Db |||||  
 286 GlyPheProGlyAlaProGlyLeuPro---GlyLeuLysGlyHisArgGlyHisLysGly 304  
 QY 143 CTG----- 145  
 Db |||||  
 305 LeuGluGlyProLysGlyGluLeuGlyAlaProGlyAlaLysGlyGluAlaGlyProThr 324  
 QY 146 GCAGAAATGGCGCTGG-----CTGATTCTAGGAGTGGCGGCGCAGGAGGAGA 199  
 Db |||||

Db 325 GlyProMetGlyAlaMetGlyProLeuGlyProArgGlyMetProGlyGluArg-GlyAr 344  
 QY 200 GGCCGAGCTTCTGGAGCAGAGCCGAGAGCAGTTCCTGGAGTCTGCTGAACGCC-- 257  
 Db |||||  
 344 gLeuGlyProGlnGlyAlaProGlyLysArgGlyAlaHisGlyMetProGlyLysProG 364  
 QY 257 ----- 257  
 Db 364 yProMetGlyProLeuGlyIleProGlySerSerGlyPheProGlyAsnProGlyMet 384  
 QY 258 -----CCTGA 262  
 Db 384 sGlyGluArgGlyProHisGlyAlaArgGlyProGluGlyProGlnGlyGlnArgGly 404  
 QY 263 GCCCTACCCGCTGGCCCACTATGCTCCAGAGGCTGT----- 299  
 Db 404 uThrGlyProProGlyProAlaGlySerGlnGlyLeuProGlyAlaValGlyThrAspG 424  
 QY 300 -----GGGTGAGCCGCTGCTGCGGCCACC---GGAAGCCCACTCTTG 340  
 Db 424 yThrProGlyArgLysGlyAlaThrGlySerAlaGlyThrSerGly-ProProGlyLeu 444  
 QY 341 CTGTGCAACCTGCTAACCTTTGGCTGAGGTGTGTTGGCCGAGGATCACCTATGTG 400  
 Db 444 laGlyProProGlySerProGlyProGlnGly----- 454  
 QY 401 CGGCTCTGCTGCTGGAAGTGGGCTAGAGAGAGTTCATGACCATGCTGCT---GGGC 457  
 Db 455 -----SerThrGlyProGlnGlyIleArgGlyGlnSerGlyAspProGlyValProG 473  
 QY 458 ATTGTCTCAGTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 517  
 Db 473 heLysGlyGluAlaGlyProLysGlyGluProGlyProHisGlyIleGlnGlyPro 493  
 QY 518 CGTGAGCAGCTATGGCCGCG-----CCGCCCTTCATCTGGGCA 556  
 Db 493 ly-----ProProGlyGluGluGlyLysArgGlyProArgGlyAspProGly 509  
 QY 557 CTGTC-----CTTGGCATCTGCTGAGCT----- 582  
 Db 509 hrValGlyProProGlyProMetGlyGluArgGlyAlaProGlyAsnArgGlyPhePro 529  
 QY 583 -----CTTTCATCCCAAGGCGGCTGGCTAGCAGGCTGCTGCTGCTGCTGCTG 631  
 Db 529 lySerAspGlyLeuProGlyProLysGlyAlaGlnGlyGluArgGlyProVal---Gly 548  
 QY 632 CCCAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688  
 Db 548 erSerGlyProLysGlyGlnGlyAspProGlyArgProGlyGluProGlyLeu--- 566  
 QY 689 GGCCAGGTGCTTCTCCTCCTGAGGCGCTGCT-----CTCT 727  
 Db 567 ---ProGlyAlaArgGlyLeuThrGlyAsnProGlyValGlnGlyProGluGlyLys 586  
 QY 728 GACCTCTCCGGACCCGGA-----CCACTGTGCGCAGGCTTACTCTGCTGCTAT 775  
 Db 586 lyProLeuGlyAlaProGlyGluAspGlyArgProGlyProGlySerIleGlyIleA 606  
 QY 776 GCTTCATGATCAGTCTTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835  
 Db 606 rgGlyGlnProGlySerMetGlyValProLysGlySerGlySerGlyAspLeuGly 626  
 QY 836 ACCAGTGGCTTGGCCCTTACCTGGGACCCAGGA-----GGAGTGGCTCTTTGGCCTG 889  
 Db 626 ys-----ProGlyGluAlaGlyAsnAlaGlyValPro----- 636  
 QY 890 CTCACCTCATCTCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949  
 Db 637 -----GlyGlnArgGlyAlaProGlyLysAspGlyGlu 648  
 QY 950 CTGGGCGCCAC-----CGACCCAGCAGAGGCTGCTGCGCC 985  
 Db 648 aGlyProSerGlyProValGlyProProGlyLeuAlaGlyGluArgGlyGluAlaGly 668









[illegible]



```

QY 172 AGGAATCAGCAGCGCCCATTTCTGCCAGCCCTTTGGT----- 134
      :::      |||      |||      |||
Db 758 GlyGluArgGlyAlaAlaGlyIleSerGlyProLysGlyAspArgGlyAspThrGlyGlu 777
QY 133 -----GCCGGTCCAGCTTCT 119
      |||      |||      |||      |||
Db 778 LysGlyProGluGlyAlaProGlyLysAspGlySerArgGlyLeuThrGlyProLeuGly 797
QY 118 CAGCCCATGCTCAACACACTGCTGCTGTGGGGCCACCTCAGTGGGGACACACTCTCATCACTC 59
      |||      |||      |||      |||      |||      |||
Db 798 ProPro-----Gly-ProAlaGlyProAsnGlyGluLysGlyGluSerGlyProSerGlu 815
      |||      |||      |||      |||      |||      |||
QY 58 AGATCTTGGC---CGAGGCGCGGGCTGCACCCGGA 25
      |||      |||      |||      |||      |||      |||
Db 815 yProProGlyIleValGlyAlaArgGlyAlaProGly 827
      |||      |||      |||      |||      |||      |||

RESULT 41
TI4339
sucrose-proton transport protein - carrot
N:Alternate names: sucrose/H+ symporter protein
C:Species: Daucus carota (carrot)
C:Title: Characterization of source- and sink-specific sucrose/H+ symporters from
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14339
R:Shakya, R.; Sturm, A.
Plant Physiol. 118, 1473-1480, 1998
A:Title: Characterization of source- and sink-specific sucrose/H+ symporters from
A:Reference number: Z17991; MUID:99063785; PMID:9847123
A:Accession: T14339
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-501 <SHA>
A:Cross-references: EMBL:Y16766; NID:g2969886; PIDN:CAA76367.1; PID:g2969887
A:Experimental source: cultivar Nantaise; leaf
C:Genetics:
A:Gene: SUT1a
C:Superfamily: common tobacco sucrose transport protein

Alignment Scores:
Pred. No.: 4.08e-11 Length: 501
Score: 309.00 Matches: 131
Percent Similarity: 37.44% Conservative: 94
Best Local Similarity: 21.80% Mismatches: 219
Query Match: 4.81% Indels: 157
DB: 2 Gaps: 16

US-09-759-143-110 (1-3410) x T14339 (1-501)

```







```

Db      675 leProGlySerLeuGlyLysCysGlyAspProGlyLeuProGlyProAspGlyGluProG 695
QY      865 GGTGCCAGGTAGGGG-----CCAG 845
Db      695 lytleProGlyIleGlyPheProGlyProGlyProGlyProGlyAspGlnGlyPheProG 715
QY      844 GGAC-----TGGTGTCACGTCATATGGCAGCAGGAGGTAGCCAGCAGCAGC 797
Db      715 lyThrLysGlySerLeuGlyCysProGlyLysMetGlyGluProGlyLeuProGlyLysP 735
QY      796 CCCAAGACTCATCAAGGCATAGACAGTAGGTAGGCTGGCCACAGTGGTCCG----- 744
Db      735 ro-----GlyLeuProGlyAlaLysGlyGluProAlaIav 746
QY      743 -----GGTCCCGGAAGA-----GGTCAGAGAGCA----- 720
Db      746 aAlaMetProGlyGlyProGlyThrProGlyPheProGlyGluArgGlyAsnSerGlyG 766
QY      719 -----GGCCCTCCAGTGAGTGAAGCACACCTCGCCAGCCTCCAGGGCGCTGGATC 629
Db      766 luHisGlyGluIleGlyLeuProGlyLeuProGlyLeuProGlyThrProGlyAsnGluG 786
QY      682 GTCCACAGCCACGC-----CCAGGATGACAGTCCAGCTCCAGGGCGCTGGATC 629
Db      786 lyLeuAspGlyProArgGlyAspProGlyGlnProGlyPro-----ProGlyGluGlnGlyP 805
QY      628 CGGGCAGCAGCCCTGTAGCAGCCGCGCTGGGATGAGAAAGAGGCTCAGCAGGAT 569
Db      805 ro-----ProGlyArgCysIleGluGlyProArgGlyAlaGlnGlyL 819
QY      568 GCCAAGGACAGTGCACATGAAGGGCGGGCGGGCCATAGCTGCACGCCAGTGGTC 509
Db      819 euPro-----GlyLeuAsnGlyLeuLysGlyGlnGlnGlyA 831
QY      508 ACTGGCTGAGCCTAGGAGCGGACACAGACCAAGGC-----CCAGCAGCTGACCAATCCCCAG 452
Db      831 rgArgGlyLysThrGlyProLysGlyAspProGlyIleProGlyLeuAspArgSerGlyP 851
QY      451 CACCATGTCATGAACCTCTCTCTACCCCGCTTCCAGCAGCAGAGCGGCGCATAGGT 392
Db      851 heProGlyGluThrGlySerProGlyIleProGlyHisGlnGlyGluMetGly----- 868
QY      391 GATGCTGCGGCGCAACACACATCCAGCCCAAGG-----TTAG 353
Db      869 -----ProLeuGlyGlnArgGlyTyPrProGlyAsnProGlyIleL 882
QY      352 CAGGTTGACCAAGCAGCTGGGCTTTCGGTGCGCGCAGCAGCGGCTCACCACAGCCT 293
Db      882 euGlyProGlyGluAspGlyValIleGlyMetMet-----GlyPheProGlyAlaI 900
QY      292 CTGGACCATAGTGGG-----CCAGGCGGTAGGGCTCAGGGGGCGGTTCAGGCATCCAGAA 236
Db      900 leGlyPro-----ProGlyProGlyAsnProGlyThrProGlyHisArgLysProGly 919
QY      235 CTGCTCTGCTCGGCTCTCTCCAGACCTGGGGCTCTCTCTCTGTCGCGGCACTG 176
Db      920 ile-----ProGlyValLysGlyGln--- 926
QY      175 CTAGGAATCAGCCAGCGGCCCATTTCTGCCAGCCTTTGTCGCGGCTCAGCTCTCTCAG 116
Db      927 ---ArgGlyThrProGlyAlaLysGlyGluGln-----GlyAspLysGlyAsn 941
QY      115 CCATGCTCAACACCTGTGCTGGGGACCTCAGTGGGACACGCTCTCATCTACTCAGA 56
Db      942 Pro-----Gly-ProSerGluIleSerHisValIleGlyAsp-----LysGlyG 956
QY      55 TCCTGSC---CCAGGCGCGCGCTCTCACCAGCGGCGGCTCCAGCGCTGGTCCC 1
Db      956 uProGlyLeuLysGlyPheAlaGlyAsnProGlyGluLysGlyAsnArgGlyValPro 975

```

RESULT 43

CGB07S

collagen alpha 1(III) chain - bovine

C:Species: Bos primigenius taurus (cattle)  
 C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 07-May-1999  
 C:Accession: A02862; A38001; A38002; A38003; A38005; S71946  
 R:Fietzek, P.P.; Allmann, H.; Rautenberg, J.; Henkel, W.; Wächter, E.; Kuehn, K.  
 Hoppe-Seyler's Z. Physiol. Chem. 360, 809-820, 1979  
 A:Title: The covalent structure of calf skin type III collagen. I. The amino acid seq  
 A:Reference number: A02862; MUID:80026026; PMID:488906  
 A:Accession: A02862  
 A:Molecule type: protein  
 A:Residues: 1-242 <FTE>  
 R:Dewes, H.; Fietzek, P.P.; Kuehn, K.  
 Hoppe-Seyler's Z. Physiol. Chem. 360, 821-832, 1979  
 A:Title: The covalent structure of calf skin type III collagen. II. The amino acid se  
 A:Reference number: A38001; MUID:80026027; PMID:488907  
 A:Accession: A38001  
 A:Molecule type: protein  
 A:Residues: 243-422 <DEWI>  
 R:Bentz, H.; Fietzek, P.P.; Kuehn, K.  
 Hoppe-Seyler's Z. Physiol. Chem. 360, 833-840, 1979  
 A:Title: The covalent structure of calf skin type III collagen. III. The amino acid s  
 A:Reference number: A38002; MUID:80026028; PMID:488908  
 A:Accession: A38002  
 A:Molecule type: protein  
 A:Residues: 423-571 <BEN>  
 R:Lang, H.; Glanville, R.W.; Fietzek, P.P.; Kuehn, K.  
 Hoppe-Seyler's Z. Physiol. Chem. 360, 841-850, 1979  
 A:Title: The covalent structure of calf skin type III collagen. IV. The amino acid se  
 A:Reference number: A38003; MUID:80026029; PMID:488909  
 A:Accession: A38003  
 A:Molecule type: protein  
 A:Residues: 572-808 <LAN>  
 R:Dewes, H.; Fietzek, P.P.; Kuehn, K.  
 Hoppe-Seyler's Z. Physiol. Chem. 360, 851-860, 1979  
 A:Title: The covalent structure of calf skin type III collagen. V. The amino acid seq  
 A:Reference number: A38004; MUID:80026030; PMID:488910  
 A:Accession: A38004  
 A:Molecule type: protein  
 A:Residues: 809-947 <DEW2>  
 R:Allmann, H.; Fietzek, P.P.; Glanville, R.W.; Kuehn, K.  
 Hoppe-Seyler's Z. Physiol. Chem. 360, 861-868, 1979  
 A:Title: The covalent structure of calf skin type III collagen. VI. The amino acid se  
 A:Reference number: A38005; MUID:80026031; PMID:488911  
 A:Accession: A38005  
 A:Molecule type: protein  
 A:Residues: 948-1049 <ALL>  
 A:Experimental source: skin  
 R:Henkel, W.  
 Biochem. J. 318, 497-503, 1996  
 A:Title: Cross-link analysis of the C-telopeptide domain from type III collagen.  
 A:Reference number: S71946; MUID:96404897; PMID:8809038  
 A:Accession: S71946  
 A:Molecule type: protein  
 A:Residues: 87-106; 1017-1029; 1037-1049 <HEN>  
 C:Comment: Prolines at the third position of the tripeptide repeating unit (G-X-Y) ar  
 C:Superfamily: The type III collagen molecule is a trimer of identical chains, linked to  
 C:Keywords: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo  
 F:1-1049/Product: collagen alpha 1(III) chain #status experimental <CAB>  
 F:1-14/Region: amino-terminal nonhelical telopeptide  
 F:15-1040/Region: helical  
 F:587-589/Region: cell attachment (R-G-D) motif  
 F:752-754/Region: cell attachment (R-G-D) motif  
 F:875-877/Region: cell attachment (R-G-D) motif  
 F:878-880/Region: cell attachment (R-G-D) motif  
 F:935-937/Region: cell attachment (R-G-D) motif  
 F:1041-1049/Region: carboxyl-terminal nonhelical telopeptide  
 F:95,107,119,338,950/Modified site: 5-hydroxylysine (Lys) #status experimental  
 F:107,950/Modified site: allysine (Lys) #status predicted  
 F:107/Binding site: carbohydrate (Lys) (covalent) #status experimental  
 F:1040,1041/Disulfide bonds: Interchain #status predicted

Alignment Scores:

Pred. No.: 4.69e-11 Length: 1049

Score:	308.00	Matches:	298
Percent Similarity:	31.0%	Conservative:	63
Best Local Similarity:	25.6%	Mismatches:	399
Query Match:	4.80%	Indels:	401
DB:	1	Gaps:	65

US-09-759-143-110 (1-3410) x CGB07S (1-1049)

26	CGGGTGACAGCCGCGCGCTCGCCAGGATCTGATGATGACAGCTGTCCCACTGAGG	85
QY		
Db		
134	ProGlyGluaspGlyAlaProGlyPro-----MetGlyProArg-Gl	147
QY		
86	TGCCCCACAGCAGAGTGTTG-----	107
QY		
147	YAlaProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsnAs	167
QY		
108	---AGCATGGCTGAGAAAGCTGGACCGGACCAAAAGGCTGGCAAAATGGGCGCTGG	163
QY		
167	pGlyAlaArgGlySerAspGlyGlnProGlyProProGly-----ProProGly	183
QY		
164	C-----TGATTCCTAGGCTAGTGGCGGACGACAAGA-----GGAGAGGC	202
QY		
183	YThrAlaGlyPheProGlySerProGlyAlaLysGlyGluValGlyProAlaGlySerPr	203
QY		
203	CGCAGCTCTGGACGAGCGGACGAGCAAGCAGTCTTGGAGTGCCTGACGCGCCCTTGA	266
QY		
203	oGlySerSerGlyAlaProGlyGlnArgGlyGluProGlyProGlnGlyHisAlaGlyAl	223
QY		
263	GCCTTACCGCTGGCCCATCTAGT-----CCAGAGGCTGTGGGTGAG---C	307
QY		
223	aProGlyProProGlyPro-ProGlySerAspGlySerProGlyGlyLysGlyGluMetG	243
QY		
308	CGCTGCTGCGGACCGGAAGCCAGCTCTT-----GCTGGTCAACTGCTCAACC	358
QY		
243	lyProAlaGlyIleProGlyAlaProGlyLeuLeuGlyAlaArgGlyProProGlyProp	263
QY		
359	TTTGGCCTGGAGGTGTGTTGGCGCAGGCAATCACCTATGTGCGGCTGTGCTGTGAA	418
QY		
263	roGlyThrAsnGlyValProGlyGlnArg-----GlyAlaAlaGlyG	277
QY		
419	GTGGGGGTAGA---GGAGAAGTTCAATGACCATGGT---GCTGGGCATTTGGTCCAGTGTGCT	472
QY		
277	luProGlyLysAsnGlyAlaLysGlyAspProGlyProArgGlyGluArgGlyGluAlaG	297
QY		
473	GG---CCTGGTCTGTGCTCCGCTCTAGGCTCAGCAGTACCATGGCGGTGGAGCGCTAT	529
QY		
297	lySerProGlyIleAla-----	302
QY		
530	GGCGCGCGCGCGCCTTCATCTGGGCACCTGCTGTCGGCATCCTGTGAGCCTCTTTCTC	589
QY		
303	-----GlyProGlyGlyGluAspGlyLysAsp---GlySerProGlyGlyGluProGlyAlaA	320
QY		
590	ATCCCAAGGCGCGCTGGCTAGCAGGCTGCTGTGCGCGGATCCCAAGCCCTCGAGGTG	649
QY		
320	snGlyLeuProGlyAlaAlaGlyGluArgGlyValProGlyPheArgGlyProAlaGlyA	340
QY		
650	GCACCTGCTCATCTCTGGG-----	666
QY		
340	laAsnGlyLeuProGlyGlyGluLysGlyProProGlyAspArgGlyGlyProGlyProAlaG	360
QY		
667	-----CGTGGGGCTGCTGGACTTCTGTGGCAGGTGTGCTTCACTCCACTGGAGGCCCTG	721
QY		
360	lyProArgGlyValAlaGly-----GluProGlyArgAsnGlyLeuProGlyGlyProG	378
QY		
722	CNCTCTGACCTCTCCGGGACCGGACCACTGTGCGGACGCTACTCTCTCTATGCTTC	781
QY		
378	lyLeuArgGlyIleProGly---SerProGlyGlyProGly-----	390
QY		
782	ATGATCAGTCTTGGGGCTGCTGGCTACTCTCTGCTGCCATTCGCTGACCTGGGACACCACT	841
QY		
391	-----SerAsnGlyLysProGlyProProGlySerGlnGlyGluThrGlyArg----	406
QY		
842	GGCCTGGGCCCCCTACTGGGACCCAGGA-----GGAGTGCCTCTTTGGC	886
QY		

407	Db	 --ProGlyPro---ProGlySerProGlyProArgGlyGlnProGlyValMetGlyPheP	425
887	Qy	CTGCT 	891
425	Db	roGlyProLysGlyAsnAspGlyAlaProGlyLysAsnGlyGluArGlyGlyProGlyG 445	445
892	Qy	-----CACCTCATCTTCCTACCTGCGTAGCAGCCACACTGCT---GGTGGCTAGGAG 943	943
445	Db	lyProGlyProGlnGlyProAlaGlyProAlaGlyLysAsnGlyGluThrGlyProGlnGlyProG 465	465
944	Qy	GCAGCGCTGGCCCCCACCAGCCAGCAGAGAGGGCTGTCGGCCCCCTTCCTGTGCGCCCCAC 1003	1003
465	Db	lyProThrGlyPro---SerGlyAsPlysGlyAspThrGlyProProGlyProGlnGlyL 484	484
1004	Qy	TCCTGTCCATGCTCGGGCCGCTGGCTTTCG---GAACCTGGCGCCCTCTTCCTCCCGG 1060	1060
484	Db	euGlnGlyLeuProGlyThrSerGlyProProGlyGluAsnGlyLysProGlyGlyProG 504	504
1061	Qy	CTGCACCACTGTGCTGCGCATGCCCCCACCCTCGCGCGCTCTTCGTGCTGCTGAGCTG 1120	1120
504	Db	ly-----ProLysGlyGluAlaGlyAlaProGlyLleGlyroGly----- 516	516
1121	Qy	TCGAGCTGATGGCACTCATGACCTTCACGCTGTT---TTACACGGAATTTCGTGGGGCAG 1177	1177
517	Db	-----GlyLysGlyAspSerGlyAlaProGlyGluArGlyProProGlyAlaG 533	533
1178	Qy	GGGCTCTACCAAGGGCTGCCAGAGCTGAGCGGCGCACGAGCGCCGCGAGACACTATGAT 1237	1237
533	Db	lyGlyProProGly---ProArGlyGlyAlaGlyProProGlyProGlyGly-----G 550	550
1238	Qy	GAAGCGTTCGGATGGCAGCCTGGG-----GCTGTCTCTGCTAGTCGCGC 1282	1282
550	Db	lyLysGlyAlaAlaGlyProProGlyProProGlySerAlaGlyThrProGlyLeuGlnG 570	570
1283	Qy	ATCTCCCTGGTCTCTC-----TCGTGTCATGGACCGCGCTGGTGCAGCATTCGGC 1333	1333
570	Db	lyMetProGlyGluArGlyGlyProGlyGlyProGlyProLysGlyAspLysGlyGlyuP 590	590
1334	Qy	ACTCGAGCAGCTCTATTGGCCAGCTGTGGCAGCTTTCCTCTGGCTGCGGTGCCACATGC 1393	1393
590	Db	roGlySerSer-----GlyValAspGlyAla-----p 599	599
1394	Qy	CTGTCCACAGTGTGCCGTGTGTGACAGCTTCACGCCCTCACCGGGTTCACCTTCTCA 1453	1453
599	Db	roGlyLysAspGlyProArGlyProThr-----G 609	609
1454	Qy	GCCTCAGATCTCGCTTCACACACTGGCTTCCTCTACCCACGGGAGACAGGCTTTC 1513	1513
609	Db	lyProIleGlyProProGlyProAlaGlyGlnProGlyAsPlysGlyGluSerGlyAlaP 629	629
1514	Qy	CTGCC-----CAAAATACCGAGGGGACACTGGAGGTCTAGCAGTAGGACAGC 1561	1561
629	Db	roGlyValProGlyLleAlaGlyProArGly-----GlyProGlyGluA 644	644
1562	Qy	CTGATACCACTTCCTGCCAGGCCCTTAAG-----CCTGGAGCTCCCTTCCCTAAATGGA 1615	1615
644	Db	rgelyGluGlnGlyPro-ProGlyProAlaGlyPheProGlyAlaProGlyGlnAsnGly 663	663
1616	Qy	CAGTGGGTGCTGGAGGCGAGTGGCCTGTCTCCA-----CCTCCACCC 1657	1657
664	Db	GluProGlyAlaLysGlyGluArGlyAlaProGlyGlyLubLysGlyGluGlyProPro 683	683
1658	Qy	GCCTCTCGGGGCCCTCGCTGTGATGTCTCGCTAGCTGTGTGGTGGTGAGCCCCACC 1717	1717
684	Db	Gly-AlaAlaGlyProAlaGlyGlySerGlyPro-----AlaGlyProPr 698	698
1718	Qy	G-----AGCCAGGGTGGTTCGGGCCCGGGGCATC 1747	1747
698	Db	oGlyProGlnGlyValLysGlyGluArGlySerPro-GlyGlyProGlyAlaAlaGlyP 718	718
1748	Qy	TGCTTGACCTCGCCATCTCCGATAGTGCCTTCTGCTGTCTCCA-----GGTGCC 1798	1798

Db 718 heProGlyGlyArgGlyPro-----ProGlyProProGlySerAsnGlyAsnPr 734  
Qy 1799 CCATCCCTGTTATGGCTCCATGTCAGCTCAGCCAGCTCTCTCACTGCTATATGGTG 1858  
Db 734 roGlyProProGlySerSerGlyAlaProGlyLysAsp-----GlyP 748  
Qy 1859 TCATGCCAGCCCTGGCTGGCGCAATTACTTTGCTACACAGTAGTATTGACAG 1918  
Db 748 roProGlyProProGlySerAsn-GlyAlaProGlySerProGlyLysProGly 767  
Qy 1919 AGCAGCTTGGCCAAATACTCAGCGTAGAAACACTTGGGGTGGAGGCGCTGC 1978  
Db 768 GlyAsp-----SerGlyProProGlyGluArgGlyAla 778  
Qy 1979 CTCAGTGGTCCAGCTCCCGCTCTGTTAGCCCATGGGCTGGCGGCGCA 2038  
Db 779 -----ProGlyProGlnGlyProProGlyAlaProGly-----Pro 790  
Qy 2039 GTTCTCTGCTGCCAAAGTAATGTGCTCTGCTGCTGCCACCTGTGCTGAGGTGG 2098  
Db 791 LeuGlyLeuAlaGly-----LeuThrGlyAlaArg----- 800  
Qy 2099 TAGCTGCACAGCTGGGGCTGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2158  
Db 801 -----GlyLeuAlaGlyProProGlyMet-ProGlyAlaArgGlySerPr 815  
Qy 2159 TGACTGGAGCCCTTCAAGGGGTTTCAGTCTGGACTTATACAGGGAGG-----CCAGA 2212  
Db 815 oGlyProGlnGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 830  
Qy 2213 AGGCTCCATGCAGTGAATGGGGGACTCTGCAGGTGATWACCAGCTCAGGTAA 2272  
Db 830 rGlyGlnAsnGlyGluArgGlyProProGlyProGlnGlyLeuProGlyLeuAlaGly 850  
Qy 2273 CAGCTACGCTCTAGTTGAGAC-----ACACCTAGAGAAGG 2308  
Db 850 rAlaGlyGluProGlyArgAspGlyAsnProGlySerAspGlyLeuProGlyArgAspGly 870  
Qy 2309 GTTTTGGAGCTGAATAAATCAG-----TCACCTGGTTTCCCATCT----- 2351  
Db 870 yAlaProGlyAlaLysGlyAspArgGlyGlyGlyGlyGlyGlyGlyGlyGly 890  
Qy 2352 -----CTAAGCCCTTAACCTCCAGCTCTGTTTATGTAG 2386  
Db 890 oGlyHisProGlyProProGlyProValGlyProAla----- 902  
Qy 2387 CTCTGCATGGAGTTTCTAGTAGTAACACTCTCTCTCTCTCTCTCTCTCTCTCTCT 2446  
Db 902 ----- 902  
Qy 2447 TTTGTAGGGAGAGCTCTGAGGGGCAACACACAGACACAGCCAGTCCCTCAGCCACAGC 2506  
Db 903 -----GlyLysSerGlyAspArgGlyGlyGlyGlyGlyGlyGlyGlyGly 916  
Qy 2507 ACTGTCTTTTGTGTATCACCACCCCTCTACCTTTATCAGAGTGTGCTGTTGGTCC 2566  
Db 917 -----SerGlyAlaProGlyProAlaGlySerArgGlyProProGlyPr 931  
Qy 2567 T-----TCTGTGTGCATCAGCAGA 2584  
Db 931 oGlnGlyProArgGlyAspLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 951  
Qy 2585 GACACAGCATTTAATATTTTAACTATTATTTTAACTAAGTAGAAGGAATCCATCTCT 2644  
Db 951 yHisArgGlyPhe-----ProGlyAsnPro----- 959  
Qy 2645 AGCTTTTCTGTGTGTGTCTAATATTTTGGGTAGGGTGGGGGATCCCAACAATCAGGTC 2704  
Db 960 -----GlyAlaPro----- 964  
Qy 2705 CCTGAGATAGTGTGCTATGTTGGCTGATCATTTGCCAGATCTCTCTCTCTCTCTCTCT 2764  
Db 964 rProGlyProAlaGlyHisGlnGly-----AlaValGlySerPr 977

Qy 2765 CCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTTCTACTCATCCCAATGATAATTCCA 2824  
Db 977 oGlyProAlaGlyProArgGlyProValGlyProSerGlyProPro----- 992  
Qy 2825 AATGCTGTTTACCACCAAGGTAGGGTGTGAAGAAGGTAGAGGGTGGGGCTTCCAGTCTCA 2884  
Db 993 -----GlyLys-AspGlyAlaSerGly----- 999  
Qy 2885 ACGGCTTCCCTAACCAACCCCTCTCTCTCTGCGCCAGCTGGTTCCTCCCACTTCCACTCC 2944  
Db 1000 -----HisProGlyProLeuGlyPro-ProGlyProArgGlyAsnArgGly 1014  
Qy 2945 CCTCTACTCTCTAGGACTGGGTGATGAAGGCACTGCCCAAAATTTCCCACTTACCCTCCA 3004  
Db 1015 GluArgGlySerGluGly-----SerProGlyHisProGlyGlnPro 1028  
Qy 3005 ACTTTCCTCTACCCCAACTTTCCCACTCCACACCCCTGT 3049  
Db 1029 GlyProProGlyPro-----ProGlyAlaProGlyProCys 1040

RESULT 44  
B40333  
collagen alpha 1(II) chain precursor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: B40333  
R:Su, M.W.; Suzuki, H.R.; Bleker, J.J.; Solursh, M.; Ramirez, F.  
J. Cell Biol. 115, 565-575, 1991  
A:Title: Expression of two nonallelic type II procollagen genes during Xenopus laevis  
A:Reference number: A40333; MUID:92011898; PMID:1918153  
A:Accession: B40333  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1486 <SUA>  
A:Cross-references: GB:M63595  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix  
F:37-96/Domain: von Willebrand factor type C repeat homology <VMC>  
F:1258-1486/Domain: fibrillar collagen carboxyl-terminal homology <CC>

Alignment Scores:  
Pred. No.: 4.69e-11 Length: 1486  
Score: 308.00 Matches: 277  
Percent Similarity: 33.50% Conservative: 54  
Best Local Similarity: 28.04% Mismatches: 362  
Query Match: 4.95% Indels: 296  
DB: 1 Gaps: 61

US-09-759-143-110 (1-3410) x B40333 (1-1486)

Qy 2569 GAAGGACCAACAGCCACATCTCTGATAAAGGTAAGAGGGGTGGATCAGCAAAAGAC 2510  
Db 130 GlnGlyProSerGlyGluGlnGlySerArgGlyGluArgGly---AspLysGlyGluLys 148  
Qy 2509 AGTCCTGTGGCTCAGGGGACCTGGTCTTGTGTGTGTCCTCAGGACTCTTCCCTAC 2450  
Db 149 GlyAlaProGly-----ProArgGly----- 155  
Qy 2449 AATAAGTCATATGTTCAAAATCCCATGAGGAGGTGTTTCATCTTAGAAGTCCCATGCAA 2390  
Db 155 ----- 155  
Qy 2389 GAGCTACATTAAAGAAAGCTGCAGGTTAAGGGCTTAGAGATGGAAACAGGTGACTGA 2330  
Db 156 -----ArgAspGlyGluProGly----- 161  
Qy 2329 GTTTATTACAGTCCCAAAACCCCTTCTCTAGGTGTCTCAACTAGGAGGCTAGCTGTTA 2270  
Db 162 -----ThrProGlyAsnProGly----- 167  
Qy 2269 ACCGTGAGCTGGGTAAATCCACCTGCAGAGTCCCGCATTCACGTGATGAGGCCCTTCT 2210



```

QY 362 -----CAAAGTTAGCAGTTGA 345
Db 821 GlyAlaProGlyAspArgGlyGluAsnGlyProProGlyProAlaGlyPheAlaGlyPro 840
QY 344 CCAGCAAGAGCTGGGCTTCCGGTCCCGCACA 312
Db 841 ProGlyAlaAspGlyGlnSerGlyLeuLysGlyAspGlnGlyGlnLysGly 860
QY 311 ---GGCGGCTCACCACAGCCTCTGGACCATAGTGGG---CCAGGCGGTAGGCTCAGG 258
Db 861 AspAlaGlyAlaProGlyProGlnGlyPro-SerGlyAlaProGlyProGlnGlyProth 880
QY 257 GGGCGGTTCCAGCACTCCAGAACTGCTTCTGGCTCTGGCTCTGCTCCAGAAAGCTCGGCCCTC 198
Db 880 rGlyValPheGly---ProLysGlyAlaArgGlyAlaGlnGlyProAlaGlyAlaThrG 899
QY 197 TCCTCTTGTCTGCGCCCACTGCTAGCAATCAGCCAGCGGCCCATTTCTGCGAGCCCTT 138
Db 899 yPhePro-GlyAlaAlaGlyArgValGlyThrProGlyProAsnGlyAsnProGlyPro 919
QY 137 TGGTCCGGTCCAGCTTCTCAGCCCATGCTCAACA-----CCTGCTGCTGTGG 90
Db 919 roGlyPro-----ProGlySerAlaGlyLysGlnGlyProLysGlyVala 934
QY 89 GGCACCTCACTGGGACACGCTCTCATCTCAGTCTGCTGCG---CCAGGCGGCGGCTGT 33
Db 934 rGlyAspAlaGlyProProGlyArgAlaGlyAspProGlyLeuGlnGlyAlaAlaGlyA 954
QY 32 CACCCGGA 25
Db 954 laProGly 956.

RESULT 45
F96741
probable sucrose transport protein F17M19.4 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: F96741
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96741
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE005173; NID:g6978914; PIDN:AAF34306.1; GSPDB:GN00141
A:Gene: F17M19.4
A:Map position: 1
C:Superfamily: Common tobacco sucrose transport protein

Alignment Scores:
Pred. No.: 5,39e-11 Length: 512
Score: 307.00 Matches: 128
Percent Similarity: 42.15% Conservative: 92
Best Local Similarity: 24.52% Mismatches: 208
Query Match: 4.78% Indels: 95
DB: 2 Gaps: 19

US-09-759-143-110 (1-3410) x F96741 (1-512)

QY 289 CCAGAGCTGTGGTGGAGCCGCTGCTCGGACCGGAGAGCCAGCTCTTGTGCTCA 348
Db 22 ProGluAspLeuGlyGlnPro-----SerPro-LeuArgLysIleIleSerValAl 38

```

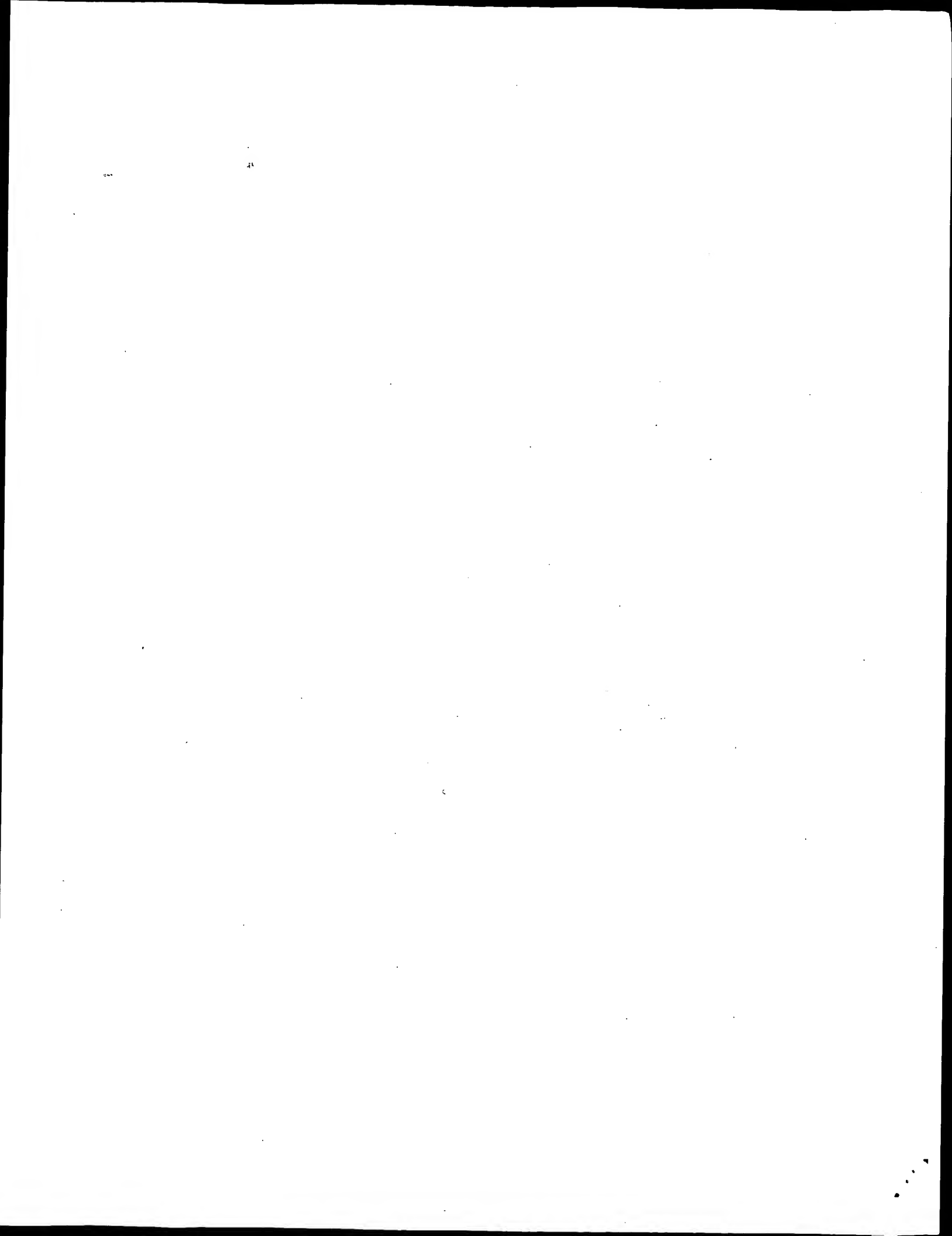
```

349 CTGCTAACCTTTGGCTGGAGGTGTGTTGGCCGAGGATCACCTATGTGCGGCTCT 408
Db 38 aserileAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeuLeuThrPro 58
QY 409 GCTGCTGAAGTGGGGTAGAGGAAGTTCATGACCATGCTGCTGGGCAATGTCAGT 468
Db 58 rileGlnLeuLeuGlyIleProHisLysTrpSerSerMettripleLeuLeuPro 78
QY 469 GCTGGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
Db 78 eserGlyMetileValGlnProIleValGlyTyHisSerAspArgCysLeuSerArg 98
QY 529 TGGCGCGCGCGCGCTTCATCTGGGCATCTGCTGCTGCTGCTGCTGCTGCTGCT 588
Db 98 eGlyArgArgProPheIleAlaAlaGlyValAlaLeuValAlaValSerValPhe 118
QY 589 CATCCCAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627
Db 118 uile-----GlyPheAlaAlaAspMetGlyHisSerPheGlyAspLysLeu 135
QY 628 GGATCCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 687
Db 135 nLysValArgThrArgAlaIleIlePheLeuThrGlyPheTrpPheLeuAspVal 155
QY 688 TGGCCAGGTGCTTCACTCCACTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTG 744
Db 155 aasnThrLeuGlnGlyProCysArgAlaPheLeuAlaAspLeuAlaAlaGlyAsp 175
QY 745 GGACCACTGTGCGCAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
Db 175 aLysLysThrArgValAlaAsnAlaCysPheSerPheMetAlaValAlaGlyAsn 195
QY 805 GGGCTAC-----CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
Db 195 uGlyTrpAlaAlaGlySerTrpThrAsnLeuHisLysMetPheProPheThr-----Me 213
QY 835 CACAGTGGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
Db 213 tThrLysAlaCysAspIleTyCysAlaAsnLeuLysThrCysPhePheLeuSerIle 233
QY 895 CTTCTATCTTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
Db 233 rLeuLeu-----LeuIleValThrPheSerSerLeu-TrpTrpValLysAspLysGln 252
QY 955 CCCCAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
Db 252 erProGlnGlyAspLys-----GluGluLys 262
QY 1015 CCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1070
Db 262 hrSerSerLeuPhePhePheGlyGlyGlyAlaValArg----- 276
QY 1071 TGTGCTGCGCATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1130
Db 277 -----HisMetLysArgProMetValMetLeuLeuIleValThrValIleAsnTrp 294
QY 1131 TGGCACTCATACCTTACAGGATTTTACAGGATTTTACAGGATTTTACAGGATTT 1190
Db 294 leAlaTrpPheProPheIleLeuTyArgTrpMetGlyArgGluValTyGlyC 314
QY 1191 GCGTGGCCAGAGCTGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1250
Db 314 ly-----AsnSerAspGlyAspGluArgSerLysLysLeuTyArgGlnGlyVal 332
QY 1251 TGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1310
Db 332 laGlyAlaLeuGlyLeuMetPheAsnSerIleLeuLeuGlyPheValSerLeu 352
QY 1311 ACCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1370
Db 352 luSerIleGlyArgLysMetGly---GlyAlaLysArgLeuTrpGlyCysValAsn 371
QY 1371 CTGTGGCTGCGGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423

```



Search completed: June 9, 2003, 22:35:30  
Job time : 309.5 secs



Result No.	Score	Query		ID	Description
		Match	Length		
1	2861	44.6	553	US-09-233-880-113	Sequence 113, App
2	2861	44.6	553	US-10-013-896-113	Sequence 113, App
3	2861	44.6	553	US-09-895-793-113	Sequence 113, App
4	2861	44.6	553	US-09-895-814-113	Sequence 113, App

US-09-759-143-110 (1-3410) x US-09-232-880-113 (1-553)

QY 284 ATGGTCCAGAGCTCTGGGTGAGCGCCTGCTGGCGCACCGAAAGCCAGCTCTTGCTG 343  
Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20  
QY 344 CTCACACCTGCTAACCTTTGGGCTGGAGGTGTGTGGCGCCGAGCATCACTATGTCCG 403  
Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyrValPro 40  
QY 404 CCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTTTCATGACCATGGTGTGGCGCATGGT 463  
Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
QY 464 CCAGTGTGGGCTGGTGTGTGCTCCGCTCCTAGGCTCAGCCAGTACCCTGGCGTGA 523  
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
QY 524 CGCTATGGCGCGCGCCCTTCATCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCTG 583  
Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
QY 584 TTTCTCATCCCAAGGGCGGCTGTGTAGCAGGGCTGTGTGGCGGATCCAGGCCCTG 643  
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
QY 644 GAGCTGGCACTGCTCATCTGGGCTGGGCTGTGTGCTGTGTGTGTGTGTGTGTGTGT 703  
Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
QY 704 ACTCCACTGGAGGGCGCTCTCTGACCTCTTCCGGGACCGCCAGCCACTGTCCCGAGGCC 763  
Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
QY 764 TACTCTGCTATGCTTCATGATCAGTCTTGGGGCTGCTGTGGGCTACCTCTGCTGCTG 823  
Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
QY 824 ATTGACTGGGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883  
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe 200  
QY 884 GGCTGTGCTACCTCATCTTCTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943  
Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
QY 944 GCAGCGCTGGCGCCACCGAGCCAGCAGAGGCTGTGGCGCCCTCTCTGCTGCTGCTGCT 1003  
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
QY 1004 TGCTGTCCATGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1063  
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
QY 1064 CACCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123  
Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
QY 1124 AGCTGTGATGCATCATGACCTTACGCTGTTTACCGATTTCTGCTGGCGAGGGGTG 1183  
Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
QY 1184 TACAGGGCTGCCAGAGCTGAGCGGCGGACCGGAGCGGCGGAGACATATGATCAAGGC 1243  
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
QY 1244 GTTCGATGGCAGCTGGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1303  
Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
QY 1304 GTCTGAGCGGCTGGTGGAGCGATTCGGCACTCCGAGCAGTCTATTGGCGCAGTGTGCA 1363

Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
QY 1364 GCTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423  
Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380  
QY 1424 TCAGCGCGCTCACCAGGTTTACCTTCTCAGCCCTGCAGCTCCTGCTGCTGCTGCTGCTG 1483  
Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
QY 1484 TCCTCTACACACCGGAGAGCAGAGTGTTCCTGCCCAATACCGAGGGACACTGGAGT 1543  
Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
QY 1544 GCTAGCAGTGGAGACAGCCTGATGACCACTTCTGCTGCTGCTGCTGCTGCTGCTG 1603  
Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
QY 1604 TTCCCTTAATGGACACGCTGCTGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1663  
Db 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProProAlaLeu 460  
QY 1664 TCGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1723  
Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
QY 1724 AGGTGCTTCCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1783  
Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
QY 1784 CTGTCCAGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1843  
Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
QY 1844 ACTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1903  
Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
QY 1904 GTAGTATTGACAGAGCGACTTGGCCAAATACTCAGCG 1942  
Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

## RESULT 2

US-10-012-896-113

; Sequence 113, Application US/10012896

; Publication No. US20020183251A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedrick, Thomas S.

; APPLICANT: Carter, Darick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; APPLICANT: Wantanabe, Yoshihiro

; APPLICANT: Meagher, Madeleine Joy

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C27  
 CURRENT APPLICATION NUMBER: US/10/012,896  
 CURRENT FILING DATE: 2001-12-10  
 NUMBER OF SEQ ID NOS: 1011  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 113  
 LENGTH: 553  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-012-896-113

Alignment Scores:  
 Pred. No.: 3 32e-151 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 9 Gaps: 0

US-09-759-143-110 (1-3410) x US-10-012-896-113 (1-553)

```

QY 284 ATGGTCAGAGGCTGGGTGAGCCGCTGCTGGCCAGCGGAAAGCCAGCTCTGTG 343
DB 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20
QY 344 GTCAACCTGTAACCTTTGGCCCTGGAGGTGTGTTGGCCGAGGATCACCTATGTGCG 403
DB 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrValPro 40
QY 404 CCTCTGCTGGAAGTGGGGTAGAGGAGAACTTCATGACCAGTGTGCTGGGCATGCT 463
DB 41 ProLeuLeuGluValGlyValGluLysPheMetThrMetValLeuGlyIleGly 60
QY 464 CCAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523
DB 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerPheHisTrpArgGly 80
QY 524 CGCTATGGCCGCGCCGCTTCATCTGGGCACTGTCTTGGGCATCTCTGCTGAGCCCTC 583
DB 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100
QY 584 TTTCTCATCCAAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
DB 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120
QY 644 GAGCTGGCACTGCTCATCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703
DB 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140
QY 704 ACTCCACTGGAGGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 763
DB 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160
QY 764 TACTCTGTCTATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823
DB 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180
QY 824 ATTGACTGGACACAGTGCCTGCGCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
DB 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200
QY 884 GGCCTGCTCACCTCATCTCTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
DB 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220
QY 944 GCAGGCTGGGCGCCACCGAGCCAGCAGAGGCTGCTGCGCCCTCTCTGCTGCGCCAC 1003
DB 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240
QY 1004 TCGCTGTCCATGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
DB 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

```

```

QY 1064 CACCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1123
DB 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280
QY 1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTACAGGATTTCTGGTGGCGAGGGCTG 1183
DB 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300
QY 1184 TACCAGGCGTGGCCAGAGCTGAGCCGGGACCGGAGCCGAGACACACTATGATGAAGGC 1243
DB 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320
QY 1244 GTTCGGATGGGACGCTGGGCTGCTCCTCAGTGCAGCTGCCCACTCCCTGGTCTTCTCTG 1303
DB 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340
QY 1304 GTCATGGACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1363
DB 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360
QY 1364 GCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
DB 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380
QY 1424 TCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1483
DB 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400
QY 1484 TCCTCTACCCACCGGAGAGGAGTGTCTGCTGCCAAATACCGAGGAGGACACTGGAGGT 1543
DB 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420
QY 1544 GCTAGCAGTAGGACAGCTGATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603
DB 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440
QY 1604 TTCCCTAATGGACAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1663
DB 441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460
QY 1664 TCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1723
DB 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480
QY 1724 AGGTGGTTCGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1783
DB 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500
QY 1784 CTGCTCCAGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1843
DB 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520
QY 1844 ACTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1903
DB 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540
QY 1904 GTAGTATTCAGAGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1942
DB 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

```

## RESULT 3

US-09-895-793-113  
 : Sequence 113, Application US/09895793  
 : Publication No. US20020192763A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Xu, Jiangchun  
 : APPLICANT: Dillion, Davin C.  
 : APPLICANT: Mitcham, Jennifer L.  
 : APPLICANT: Harlocker, Susan L.  
 : APPLICANT: Jiang, Yuqiu  
 : APPLICANT: Kalos, Michael D.  
 : APPLICANT: Retter, Marc W.  
 : APPLICANT: Stolk, John A.

```

; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895.793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-895-793-113

Alignment Scores:
Pred. No.: 3,32e-151 Length: 553
Score: 2861.00 Matches: 553
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.58% Indels: 0
DB: 9 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-895-793-113 (1-553)

QY 284 ATGGTCCAGAGGCTGTGGGTGAGCGCGCTGTGGCGACCGAAGCCAGCCTGTGCG 343
DB 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20

QY 344 GTCAACCTGTAACTTTGGCGTGGAGGTGTGTTGGCGCCAGCATCACTATGCGCG 403
DB 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyrValPro 40

QY 404 CCTCTGCTGTGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGTGTGGCGATTC 463
DB 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGly 60

QY 464 CCAGTGTGGGCGTGTCTGTGTCGGCTCCTAGGCTCAGCCAGTCACTGGCGTGA 523
DB 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

QY 524 CGCTATGGCGCGCGCGCTTCATCTGGGCACTGTCTGGGCATCTGTGGCATCT 583
DB 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

QY 584 TTCTCATCCAAAGCGCGCTGGCTAGCAGGCGTCTGTGGCGGATCCAGGCGCGCT 643
DB 101 PheLeuIleProArgAlaGlyTyrPheAlaGlyLeuLeuCysProAspProArgProLeu 120

QY 644 GAGCTGGCACTGCTCATCTGGCGGTGGGCTGTGGACTTCTGTGGCCAGGTGCTTC 703
DB 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

QY 704 ACTCCACTGGAGGCGCTGTCTCTGACCTTCCGGGACCGCCAGGACCTGTGGCG 763
DB 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

QY 764 TACTCTGTCTATGCCCTTCATGATCAGTCTTGGGGGCTGTGGGTACCTCTGCTGCC 823
DB 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

```

```

QY 824 ATTGACTGGGACACACAGTGCCTGGCCCTACTGTGGGACCCAGAGGAGTGCCTCTTT 883
DB 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200

QY 884 GGCCTGCTCACCTCATCTTCCTCAGCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
DB 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

QY 944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAGGCTCTGGGCCCCCTCTGTGCGCCCCAC 1003
DB 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

QY 1004 TGCTGTCCATGCGCGCGCTTGGCTTCCGGAACCTGGCGCCCTGCTTCCCGGCTG 1063
DB 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

QY 1064 CACCAGCTGTGTCGCGCATGCGCCACCCCTGCGCGCTCTTCTGTGCTGAGTGTGC 1123
DB 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgPheValAlaGluLeuCys 280

QY 1124 AGCTGGATGGCACTCATGACCTTCAGCGCTTTTACACCGATTTCTGTGGCGGAGGCTG 1183
DB 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlyLeu 300

QY 1184 TACAGGCGCTGCCAGAGCTGAGCGGCGCAGCGCCGAGAGCAGACTATGATGAAGC 1243
DB 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgHisTyrAspGluGly 320

QY 1244 GTTCGATGGGAGCCTGGGCTGTCTGTCAGTGGCGCATCTCCTGGTCTCTCTCTG 1303
DB 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

QY 1304 GTCATGGACCGCTGTGTCAGCGATTTCGCGACTCCAGCAGTCTATTTCGCGAGTGGCA 1363
DB 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

QY 1364 GCTTTCCTGTCGTCGCGCTGCCACATGCCTGTGCCAGTGTGGCGCTGGTGACAGCT 1423
DB 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380

QY 1424 TCAGCGCGCTCACCGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACACTGGCC 1483
DB 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

QY 1484 TCCCTCTACCGCGGAGAGAGAGGTTCCTGCCCAAATACCGAGGCGACACTGGAGGT 1543
DB 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

QY 1544 GCTAGCAGTGGAGCAGCCTGATGACCATCTCTGCCAGGCTCTGCCAGGCGCTTAAGCCTGGAGCTCC 1603
DB 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

QY 1604 TTCCCTAATGACAGCGTGGGTGCTGGAGCAGTGGCTCTGCCACCTCCACCCCGCGCTC 1663
DB 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeu 460

QY 1664 TCGCGGCGCTGCTGCTGATGCTCTCGTACGTGTGGTGGTGGTGGAGCCCGAGGCGC 1723
DB 461 CysGlyAlaSerAlaCysAspValSerValArgValValGlyGluProThrGluAla 480

QY 1724 AGGTGGTGTTCGCGCGCGCGGCGATCTGCTGGACCTCGCCATCTCCGATAGTGCCTTCTG 1783
DB 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

QY 1784 CTGTGCCAGGTGGCCCCCTCCTCTTTATGGCTCCATTGTTCAGCTCAGCAGTCTGTGTC 1843
DB 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

QY 1844 ACTGCTATATGCTGTCTGCGCGAGGCTGGGTCTGCTGCCCATTTACTTGTGTACACAG 1903
DB 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

QY 1904 GTAGTATTGTACAAGCGGAGCTTGGCCAAATACTCTACGG 1942

```



Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553  
|||||

## RESULT 4

US-09-895-814-113

; Sequence 113, Application US/09895814

; Publication No. US20020193296A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedrick, Thomas S.

; APPLICANT: Carter, Barrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C26

; CURRENT APPLICATION NUMBER: US/09/895,814

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 990

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 113

; LENGTH: 553

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-895-814-113

## Alignment Scores:

Pred. No.:	3.32e-151	Length:	553
Score:	2861.00	Matches:	553
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	44.58%	Indels:	0
DB:	9	Gaps:	0

US-09-759-143-110 (1-3410) x US-09-895-814-113 (1-553)

QY	284	ATGTCACAGAGCTGTGGGTGAGCGCGCTGCTGGCGCACCGAAGCCAGCTCTTGGT	343
Db	1	MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu	20
QY	344	GTCAACCTGTCAACCTTTGGCTGGAGGTGTGTGGCGCAGGCATCATGTGCGG	403
Db	21	ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro	40
QY	404	CCTCTGCTGCTGGAAGTGGGGTAGAGGAGGATTCATGACCATGGTGGCGCATTTGGT	463
Db	41	ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly	60
QY	464	CCAGTGTGGCGCTGTGTCTGCTCCCGCTCTAGCTCAGCCAGTGCAGCTGCGTGA	523
Db	61	ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly	80
QY	524	CGCTATGGCGCGCGCGCTTCATCTGGGCACTGTCTTGGGCACTCTCTGAGGCTC	583
Db	81	ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu	100

QY	584	TTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGCTGTGCCCGATCCAGGCGCCCTG	643
Db	101	PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspArgProLeu	120
QY	644	GAGCTGGCACTGCTCATCTGGGCTGGGCTGGACTTCTGTGGCCAGGTGTGCTTC	703
Db	121	GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe	140
QY	704	ACTCCACTGGAGCGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTGCCAGGCC	763
Db	141	ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla	160
QY	764	TACTCTGTCTATGCTTCATGATCAGTCTTGGGGCTGCTGGGTACCTCTCTGCTGCC	823
Db	161	TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla	180
QY	824	ATTGACTGGACACCACTGCTGCCCTGCTGGGACCCAGGAGGAGTCTCTTT	883
Db	181	IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe	200
QY	884	GGCTGCTCACCTCATCTTCTCAGCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAG	943
Db	201	GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu	220
QY	944	GCAGCGCTGGGCGCCACAGCAGCAGAGGCTGTGCGGCCCTCTCTGTGCTGCCCCAC	1003
Db	221	AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis	240
QY	1004	TGCTGTCCATGGCGGCGCTGTGGCTTCCGGACCTGGCGGCTGCTCCCGGCTG	1063
Db	241	CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu	260
QY	1064	CACAGCTGTGCTGGCGCATGCGCCGACCTGCGCGGCTCTTCTGGCTGAGTGTGC	1123
Db	261	HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys	280
QY	1124	AGCTGATGCACTCATGACCTTACGCTGTTTACAGGATTTCTGGGCGAGGGCTG	1183
Db	281	SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluLeu	300
QY	1184	TACAGGGGTGCGCAGAGCTGAGCGGCGCACCGAGCGCCGAGACACTATGATGAAGC	1243
Db	301	TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly	320
QY	1244	GTTCCGATGGCAGCCTGGGCTGTTCTCTGCACTGCGCCATCTCCCTGGTCTCTCTCTG	1303
Db	321	ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu	340
QY	1304	GTCATGGACCGGCTGGTGCAGCGATTTCGGCACTCGAGCAGTCTATTGGCCAGTGGCA	1363
Db	341	ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla	360
QY	1364	GCTTTCCTGCTGCTGCGGTGCGCATGCTGCTCCACAGTGTGGCGCTGGTGACAGCT	1423
Db	361	AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla	380
QY	1424	TCAGCGCGCTCCACCGGTTACCTTCTCAGCGCTCGAGATCTGCTTACACACTGGCC	1483
Db	381	SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla	400
QY	1484	TCCCTTACACCGGAGAGCAGGTGTCTTCCCAATACCGAGGGACACTGGAGGT	1543
Db	401	SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly	420
QY	1544	GCTAGCAGTGCAGACAGCTGATGACAGCTTCTGCCAGGCGCTTAAGCTGGAGCTCC	1603
Db	421	AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro	440
QY	1604	TTCCCTTAATGGACACGCTGGGTGCTGGAGGAGGTGCTGCTCCACCTCCACCGGCTC	1663
Db	441	PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu	460
QY	1664	TGCGGGGCTCTGCTGCTGTGATGTCTCCGTACGTGTGTGTGTGTGTGTGTGTGTGTGT	1723

461	CysGlyAlaSerAlaCysAspValSerValArgValValGlyGluProThrGluAla	480
1724	AGGTGGTTCGGGCGCGGGGCATCTGCCTGGACCTGCCATCCTGGATAGTCCTTCTGT	1783
481	ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu	500
1784	CTGTCCCAAGTGGCGCCCATCCCTGTTATTGGGCTCATTTGCCAGCTCAGCCAGCTCTGTC	1843
501	LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal	520
1844	ACTGCCTATATGGTGTCTGCCGAGCGCTGGGTCTGGTGGCCATTTACTTTGCTACACAG	1903
521	ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln	540
1904	GTAGTATTGCACAGCGCACTTGGCCAAATACTCAGCG	1942
541	ValValPheAspLysSerAspLeuAlaLysTyrSerAla	553

## RESULT 5

US-10-010-940-113  
; Sequence 113, Application US/10010940  
; Publication No. US20030088062A1

; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan Louise  
 ; APPLICANT: Jiang Yuqi  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Kalos, Michael  
 ; APPLICANT: Fanger, Gary  
 ; APPLICANT: Retter, Mark  
 ; APPLICANT: Solk, John  
 ; APPLICANT: Day, Craig

Alignment Scores:	
Pred. No.:	3.32e-151
Score:	2861.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	44.58%
DB:	9
Length:	553
Matches:	553
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-759-143-110 (1-3410) x US-10-010-940-113 (1-553)

284	ATGTCTCAGAGCGTGTGGGTGACGCCCTGCTCGGCACCGGAAAGCCACGCTTGTGTG	343
QY		
Db		
1	MetValGlnArgLeuTyrValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu	20
344	GTCACCTGCTAACCTTTGGCTTGGAGGTGTGTTTGGCGCAGGCATCACCTATGTGGCG	403
QY		
Db		
21	ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro	40
404	CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTTCATGACCATGGTGTGGGCATTGGT	463
QY		
Db		
41	ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly	60
464	CCAGTGTGGGCGCTGGTCTGTCTGGCTCTCCTAGGCTCAGCCAGTACACATGGCGGTGGA	523
QY		
Db		
61	ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTyrPargGly	80





1604	QY	TTCCCTAATGACACGATGGGGTCTGGAGCATGGCGTCTCCCACTCCACCCGCGCTC	1663
1605	QY		
1606	QY		
1607	QY		
1608	QY		
1609	QY		
1610	QY		
1611	QY		
1612	QY		
1613	QY		
1614	QY		
1615	QY		
1616	QY		
1617	QY		
1618	QY		
1619	QY		
1620	QY		
1621	QY		
1622	QY		
1623	QY		
1624	QY		
1625	QY		
1626	QY		
1627	QY		
1628	QY		
1629	QY		
1630	QY		
1631	QY		
1632	QY		
1633	QY		
1634	QY		
1635	QY		
1636	QY		
1637	QY		
1638	QY		
1639	QY		
1640	QY		
1641	QY		
1642	QY		
1643	QY		
1644	QY		
1645	QY		
1646	QY		
1647	QY		
1648	QY		
1649	QY		
1650	QY		
1651	QY		
1652	QY		
1653	QY		
1654	QY		
1655	QY		
1656	QY		
1657	QY		
1658	QY		
1659	QY		
1660	QY		
1661	QY		
1662	QY		
1663	QY		
1664	QY		
1665	QY		
1666	QY		
1667	QY		
1668	QY		
1669	QY		
1670	QY		
1671	QY		
1672	QY		
1673	QY		
1674	QY		
1675	QY		
1676	QY		
1677	QY		
1678	QY		
1679	QY		
1680	QY		
1681	QY		
1682	QY		
1683	QY		
1684	QY		
1685	QY		
1686	QY		
1687	QY		
1688	QY		
1689	QY		
1690	QY		
1691	QY		
1692	QY		
1693	QY		
1694	QY		
1695	QY		
1696	QY		
1697	QY		
1698	QY		
1699	QY		
1700	QY		
1701	QY		
1702	QY		
1703	QY		
1704	QY		
1705	QY		
1706	QY		
1707	QY		
1708	QY		
1709	QY		
1710	QY		
1711	QY		
1712	QY		
1713	QY		
1714	QY		
1715	QY		
1716	QY		
1717	QY		
1718	QY		
1719	QY		
1720	QY		
1721	QY		
1722	QY		
1723	QY		
1724	QY		
1725	QY		
1726	QY		

## RESULTS

```

RES001 8
US-09-759-143-113
; Sequence 113, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-759-143-113

```

Alignment Scores:	
Pred. No.:	3.32e-151
Score:	2861.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	44.58%
DB:	10
Length:	553
Matches:	553
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-759-143-110 (1-3410) x US-09-759-143-113 (1-553)

Qy		284	ATGGTCCAGAGGCTGTGGGTGAGCGGCCTGCTGGGCACCGGAAGCCCCAGCTCTTGCTG	343
Dd		1	MetValGlnArgLeuTrpValSerArgLeuMetHisArgLysAlaGlnLeuLeuLeu	20

344	QY	GTCAACCTGCTAAACCTTTGGCCCTGGAGGTGTGTTTGGCCGCGAGGCATCACTATGTGCCG	403
21	Db	ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro	40
		ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro	40
404	QY	CCTCTGCTGCTGAAGTGGGGTAGAGGAGAAGTTTATGACCATGTGTCTGGCCATTTGGT	463
41	Db	ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly	60
		ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly	60
464	QY	CCAGTGTGCGGCTGTGCTGTGTCGCCGCTCTAGGCTCAGCCAGTCACCACTGGCGTGA	523
61	Db	ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly	80
		ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly	80
524	QY	CGCTATGGCGCGCGCGCCCTTCATCTGTGGCACTGTCTTGGGCATCTGTGTGAGCCTC	583
81	Db	ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu	100
		ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu	100
584	QY	TTTCTCATCCAAAGCGCGCTGCTGTGTGTCAGCAGGCTGTGTGCCGGATCCACGCGCCCTG	643
101	Db	PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu	120
		PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu	120
644	QY	GAGCTGGCACTGTCTCATCTGTGGCGTGGGCTGCTGTGACTTCTGTGCCAGGTGTGCTTC	703
121	Db	GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe	140
		GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe	140
704	QY	ACTCCACTGAGGCCCTGTCTCTGTGACTTTCGGGACCCGGACCACTGTGTGCCAGGCC	763
141	Db	ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla	160
		ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla	160
764	QY	TACTCTCTCATGCCCTTCATGATCAGTCTTGTGGGGCTGCTGTGGCTACCTCTGCCTGCC	823
161	Db	TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla	180
		TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla	180
824	QY	ATTGACTGGGACACCACTGGCCCTGGCCCTACTTGGGCACCCAGGAGGAGTGCCTCTT	883
181	Db	IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe	200
		IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe	200
884	QY	GGCCTGTCAACCTCATCTTCCTTCACCTGGGTAGCAGCCACACTGTGTGGTGGCTAGGAG	943
201	Db	GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu	220
		GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu	220
944	QY	GCAGCGTGGGCCCCACCGAGCCAGCAGAGGCGTGTGCCGCCCTCTTGTTCGCCCCAC	1003
221	Db	AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis	240
		AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis	240
1004	QY	TGCTGTTCATGCGCGGCCCTTGGCTTTCCGAACTGGCGGCCCTGTCTTCCCCGGCTG	1063
241	Db	CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu	260
		CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu	260
1064	QY	CACCACTGTCTGCCGCATGCCCGCCACCTCGCGCGGCTCTTCTGTGGCTGAGCTGTGC	1123
261	Db	HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys	280
		HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys	280
1124	QY	AGCTGGATGCACCTCATGAOCTTACACGCTGTTTTACACGATTTCTGGGCGAGGGCTG	1183
281	Db	SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu	300
		SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu	300
1184	QY	TACCAGGGCGTGCCAGAGCTGAGCCGGGACCCAGGCGCGGAGACACTATGATGAAGGC	1243
301	Db	TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly	320
		TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly	320
1244	QY	GTTCCGGATGGCGAGCCTGGGGCTTCTCCTGCAGTGGCCCATCTCCCTGTGCTCTCTCTG	1303
321	Db	ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu	340
		ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu	340
1304	QY	GTCATGACCGGCTGGTGCAGCGATTTCGGCACTCGACAGTCTATTTGGCCAGTGTGCA	1363
341	Db	ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla	360
		ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla	360
1364	QY	GCTTTCCTGTGGCTGCGGTGGCCATGCTGTGCCACAGTGTGGCCGTGGTGCACAGCT	1423
361	Db	AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla	380
		AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla	380
1424	QY	TCAGCCGCGCCTACCGGGTTACCTCTCAGCCCTGCAGATCTCCTACACACTGGCC	1483

		I I	
381	Db	SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla	400
		TTCCTCTACACCGGAGAACGAGGTTCCTGCCAAATACGAGGGACACTGGAGGT	1543
401	QY	SerLeuTyHisargGluLysGlnValPheLeuProLysTyrargGlyAspThrGlygly	420
		GCTAGCAGTGAGACAGCCCTGATGACCAGCTTCCTGCACGGCCCTAAGCCTGGAGCTCCC	1603
421	Db	AlaserSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro	440
		TTCCTTAATTGGACACCTGGGTGCTGGAGGAGTHGGCCTGCTCCCACTCCACCCGGCCTC	1663
441	Db	PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu	460
		TGCGGGCCCTCTGCCFTGATGTCCTCCGACGTGTGGTGGGTGGGTGAGCCCCACCGAGGCC	1723
461	Db	CysGlyAlaSerAlacysAspValSerValArgvalValvalGlyGlnProThrGluAla	480
		AGGTTGGTTCCGGCCGGGGCATCTCCCTGGACCTCGCCATCTCGTATAGTGCCTTCCTG	1783
481	Db	ArgvalValproGlyArgGlylleCysLeuAspLeuAlalaileLeuAspSerAlaPheLeu	500
		CTGTCCACAGTGGCCCATCCCTGTTTTATGGGTCCATTGTCACAGTCACGAGTCTGTC	1843
501	Db	LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal	520
		ACTGCCTATATGTGTCCTCCGACGCCCTGGTCTGGTCGCATTTCATTTCCTCACACAG	1903
521	Db	ThrAlaTyMetValSerAlaLaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln	540
		GTAGTATTTACAAGAGCGACTTGGCCAAATACTACGCG	1942
541	Db	ValValPheAspLysSerAspLeuAlaLysTyrSerAla	553

## RESULT 9

[illegible]



1244 GTTCGATGGGCGAGCTGGGGCTGTCTCTGAGTGGCCATCTCCCTGGTCTCTCTCTG 1303  
Db ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
1304 GTTCGATGGGCGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 1363  
Db ValMetArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
1364 GTTCCTCTCTGGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 1423  
Db AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380  
1424 TCAGCGCGCTCACGGGTTACCTTCACGCTTCAGCGCTTCAGCGCTTCAGCGCTTCAGCGCT 1483  
Db SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
1484 TCCTCTACACCGGAGAGAGGTGTCTCTGCTCCCAATACCGAGGGGACACTGGAGGT 1543  
Db SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
1544 GCTAGCAGTGGAGACAGCTGTATGACCTTCCTGCGCAGGCGCTTAAGCTGGAGCTGCC 1603  
Db AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
1604 TTCCTTAATGGACAGCTGGTGTGGAGGACGTGGCTGCTCCACCTCCACCGCGGCTC 1663  
Db PheProAsnGlnHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 460  
1664 TGGCGGCGCTCTGCTGTGATGCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1723  
Db CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
1724 AGGTGTTCGGCGGGCGGCGATCTCCCTGGACCTCGCATCTGGATAGTGGCTTCCTG 1783  
Db ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
1784 CTGTCCAGGTGGCGCCATCTCTGTTATGGGTCCATTTGCTCCAGCTACCGCAGTGTGC 1843  
Db LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
1844 ACTGCCATATGTGTCTGCGGAGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1903  
Db ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
1904 GTAGTATTCACAGAGCGACTTGGCCAAATACCTACGCG 1942  
Db ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

## RESULT 10

US-09-030-606-113  
; Sequence 113, Application US/09030606  
; Patent No. US20020081580A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillion, David C.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS  
; NUMBER OF SEQUENCES: 224  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/030,606  
; FILING DATE: 25-FEB-1998

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.428C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 113:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-09-030-606-113

Alignment Scores:  
Pred. No.: 3,32e-151 Length: 553  
Score: 2861.00 Matches: 553  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 44.58% Indels: 0  
DB: 10 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-030-606-113 (1-553)

284 ATGTCTCAGAGGCTGTGGGTGAGCGCTGTCTGGGACCGAAGCCAGCTCTCTGCTG 343  
QY  
Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
344 CTCACCTGCTAACTTTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTG 403  
QY  
Db 21 ValAsnLeuLeuThrPheGlyLeuValValCysLeuAlaAlaGlyIleThrTyrValPro 40  
404 CCTCTGCTGTGGAAGTGGGGTGGAGAGAGTTCATGACCATGTGGCTGTGGCTGTGGCT 463  
QY  
Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
464 CCAGTGTGTGGCTGT 523  
QY  
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
524 CGCTATGGCGCGCGCGCGCTTCATCTGGGCACTGTCTTGGGCACTGTCTTGGGCACTC 583  
QY  
Db 81 ArgTyrGlyArgArgPropheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
584 TTTCATCTCCCAAGGCGCGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTG 643  
QY  
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
644 GAGCTGGCACTGTCTGT 703  
QY  
Db 121 GluLeuAlaLeuLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
704 ACTCCACTGTGGGCGCGCTGTCTCTGACCTTTCCGGGACCGCCAGCACTGTGCCAGGCC 763  
QY  
Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
764 TACTGTGTATGCGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 823  
QY  
Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyCysLeuGlyTyrLeuLeuProAla 180  
824 ATTGCTGGGACACCACTGCTGTGGCGCGCTGTGGCGCGCTGTGGCGCGCTGTGGCGCG 883  
QY  
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200  
884 GGCTGTCTCACCTTCATCTCTACCTCGTAGCAGCCACACTGTGTGTGTGTGTGTGTGTGT 943  
QY  
Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

```
QY 944 GCAGCGCTGGCCGCCAGCCAGCAGAGCGCTGTGGGCCCTCTCTGTGCGCCAC 1003
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240
QY 1004 TGCTGTCCATCGCGCGCGCTGTGGCTTCCGGAACCTGGCGCCCTCTCCCGCGGTG 1063
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260
QY 1064 CACCAGCTGTGCTGGCGATGCCCGCCAGCCCTGTGCGCGCTCTTGTGGCTGAGCTGTC 1123
Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280
QY 1124 AGCTGATGGCACTCATGACCTTACGCTGTTTACACGATTTCTGGCGGAGGCGCTG 1183
Db 281 SerTyrMetAlaLeuMetThrPheThrLeuPheThrThrAspPheValGlyGlyLeu 300
QY 1184 TACCAGGCGTGGCCAGAGCTGAGCGCGGCGCCAGGCGCCGAGACACTATGATGAAGGC 1243
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320
QY 1244 GTTCGGATGGCAGCGCTGGGCTGTCTGTCAGTGGCGCATCTCCCTGGTCTCTCTCTG 1303
Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaLeuSerLeuValPheSerLeu 340
QY 1304 GTCATGGACCGGCTGGTGCACGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGGCA 1363
Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360
QY 1364 GCTTTCCTGCTGGCTGGCGTCCACATGCTGTCCACAGTGGCGCGTGGTGCACAGCT 1423
Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380
QY 1424 TCAGCCCGCTCACCGGCTTCCACCTTCTCAGCCCTGCGAGTCTGCTCCCTACACAGTGGC 1483
Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnLeuLeuProTyrThrLeuAla 400
QY 1484 TCCCTCTACACCGGAGAGCAGGTGTTCTGCGCCAAATACGAGGGGACACTGGAGGT 1543
Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420
QY 1544 GCTAGCAGTGGAGCAGCCTGATCACCAGCTTCTGCGAGGCTTAAAGCTGAGCTCC 1603
Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440
QY 1604 TTCCTTAATGGACAGCTGGTGTGGAGCAGTGGCTGCTCCACCTCCACCGCGGCTC 1663
Db 441 PheProAsnGlyHisValGlyAlaGlyAlaGlySerGlyLeuLeuProProProAlaLeu 460
QY 1664 TGGCGGCGCTCTGCTGTGATGCTCGTACGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1723
Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480
QY 1724 AGGTGTTCGGCGCGGCGCATCTGCTGGACCTCGCCATCTGATAGTGGCTTCCCTG 1783
Db 481 ArgValValProGlyArgGlyLeuCysLeuAspLeuAlaLeuAspSerAlaPheLeu 500
QY 1784 CTGTCCAGTGGCGCCATCCCTGTTATGGCTCCATGTCAGCTCAGCCAGTGTGTC 1843
Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520
QY 1844 ACTGCTATATGTTGCTGCGCGAGCGCTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1903
Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540
QY 1904 GTAGTATTTGACAGAGCGACTTGGCCAAATACTCAGCG 1942
Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

## RESULT 11

US-09-822-827-113

: Sequence 113, Application US/09822827

: Patent No. US20020081680A1

: GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.534C1

CURRENT APPLICATION NUMBER: US/09/822,827  
CURRENT FILING DATE: 2001-03-28  
NUMBER OF SEQ ID NOS: 982

SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 113  
LENGTH: 553

TYPE: PR

ORGANISM: Homo sapien

US-09-822-827-113

Alignment Scores:

Pred. No.: 3 32e-151 Length: 553

Score: 281.00 Matches: 553

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 44.58% Indels: 0

DB: 10 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-822-827-113 (1-553)

```
QY 284 ATGGTCCAGAGGCTGGGTGAGCCCGCTGTCGCGCACCCGAAAGCCAGCTCTTGTG 343
Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20
QY 344 GTCAACCTCTAACCTTTGSCCTGGAGGTGTGTTTGGCGCAGGATCACCTATGTGCG 403
Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40
QY 404 CTCTGCTGCTGGAAGTGGGGTAGAGAGAAAGTTTCATGACCATGCTGTGGCATTTGGT 463
Db 41 ProLeuLeuLeuGluValGlyValGluLysPheMetThrMetValLeuGlyIleGly 60
QY 464 CCAGTGTGGGCTGGTGTGTCCTCCGCTCTAGGCTCAGCAGTGCACACTGGCGTGA 523
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80
QY 524 CGCTATGGCGCGCGCGCTTTCATCTGGGCACTGTCTTGGGCATCTCTGTAGCCTC 583
Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100
QY 584 TTTCTCATCCCAAGCGCGGTGGCTAGCAGGTGTGTGCCGGATGCCAGGCCCTG 643
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120
QY 644 GAGCTGGCAGCTCTCATCTGCGGCTGGGCTGTGGACTTCTGTGGCAGGTGTGCTTC 703
Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140
QY 704 ACTCCAGTGGAGCGCTCTCTGACCTCTTCGCGGACCCGACCACTGTGCGCAGGCC 763
Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160
QY 764 TACTGTGCTATGCTTCATCATCATCAGTCTGGGGCTGCTGGGCTACCTCCTCCTGCC 823
Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180
QY 824 ATTGACTGGGACACAGTGGCCCTGCGCCCTTACCTGGGCACCCAGGAGGAGTGTCTTT 883
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe 200
QY 884 GGCTGTCTCAGCTCTCTCTCCTCAGCTGCGTAGCAGCACACTGTGTGGTGGTGGAG 943
Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220
QY 944 GCAGCGCTGGCGCCCGCCAGCAGAGGCTGTGCGGCCCTCTCTTGTGCGCCAC 1003
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240
QY 1004 TGCTGTCCATCGCGCGCGCTTGGCTTTCGGAACCTGGCGCGCTGTCTTCCCGCGCTG 1063
```

```
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260
QY 1064 CACAGCTGTGCTCCGATCCCGCCGACCTGCGCGCTCTTCGTGGCTGAGCTGCG 1123
Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280
QY 1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTACCGGATTTCTGGGGAGGGCTG 1183
Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheThrAspPheValGlyGluGlyLeu 300
QY 1184 TACAGGCGGTGCCAGAGCTGAGCGGGCCAGCGCCGAGGACACTATGATGAGGC 1243
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320
QY 1244 GTTCGGATGGCAGCTGGGCTGTCTCTGAGTGGCCATCTCCCTGGTCTCTCTCTG 1303
Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340
QY 1304 GTCATGGACCGGCTGTCAGCGATTTCGGCACTCGAGCAGTCTATTTGGCCAGTGGCA 1363
Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360
QY 1364 GCTTTCCTGTGGCTGGCGGTGCCACATGCTGCCACATGTCGCGGTGGTGACAGCT 1423
Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380
QY 1424 TCAGCGCGCTCACCGGTTCCCTCTCAGCCCTGACAGTCTGCTCCCTACACACTGGCC 1483
Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400
QY 1484 TCCCTCTACACCGGAGAGCAGGTGTTCTCTGCCAAATACCGAGGGGACACTGGAGT 1543
Db 401 SerLeuTyrHisArgGlyGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420
QY 1544 GCTACAGTGGAGCAGCTGATGACAGCTTCTGCGCAGGCTTAAGCCTGAGAGCTCCC 1603
Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440
QY 1604 TTCCCTATGAGCAGCTGGGTGCTGGAGCAGTGGCTTCCACCTCCACCGCGCTC 1663
Db 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeu 460
QY 1664 TCGGGGGCTCTGCTGTGATGCTCCGTACGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1723
Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480
QY 1724 AGGTGGTTCGGGGCGGGGATCTGCTGGACCTGCGCATCTGGATGAGTGGCTTCCTG 1783
Db 481 ArgValValProGlyValGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500
QY 1784 CTGTCCAGGTGGCGCCATCCCTGTTATGGCTCCATTGTCCAGCTCAGCAGTCTGTC 1843
Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520
QY 1844 ACTGCTATATGCTGTCTGCCAGGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1903
Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540
QY 1904 GTAGTATTGACAAGCGGACTTGGCCAAATFACTCAGCG 1942
Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

## RESULT 12

```
US-09-115-453-113
; Sequence 113, Application US/09115453B
; Patent No. US20020090372A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C4
```

```
; CURRENT APPLICATION NUMBER: US/09/115,453B
; CURRENT FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-115-453-113

Alignment Scores:
Pred. No.: 3,32e-151 Length: 553
Score: 2861.00 Matches: 553
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.58% Indels: 0
DB: 10 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-115-453-113 (1-553)
QY 284 ATGTCTCAGAGGCTGTGGGTGAGCGGCTGTCTGGCGCACCGAAAGCCAGCTCTTGGCTG 343
Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20
QY 344 GTCAACCTGTAACCTTTGGCTGGAGGTGTGTTGGCCGACGATCCTATGATGCGG 403
Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40
QY 404 CCTCTGCTGTGGAAGTGGGGGTAGAGGAGAACTTATGACCATGCTGTGGCGCATGGT 463
Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60
QY 464 CCAGTGTGGGCTGTGCTGTCTCCGCTCCTAGGCTCAGCCAGTACCATCTGGCTGGA 523
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80
QY 524 CGTATGCGCGCGCGCGCTTCTATCTGGGCTGTCTGGGCTGTCTGGGCTGTCTGGCTG 583
Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100
QY 584 TTTCTCATCCCAAGGCGCGCTGTAGCAGGCTGTGTGCGGATCCGATCCAGCCGCTG 643
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120
QY 644 GAGTGTGCTGTCTATCTCTGCGGCTGTGTGGCTGTGTGGACTTCTGTGGCCAGGTGTGCT 703
Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140
QY 704 ACTCCACTGAGGCGCTGTCTGTGACTCTTCCGGGACCGGACCATCTGTCCGCGAGGCC 763
Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160
QY 764 TACTCTCTATGCTTTCATGATCAGTCTTGGGGGCTGTGTGGGCTACCTCTGCTGCTGCC 823
Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180
QY 824 ATTGACTGGGACACAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 883
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe 200
QY 884 GGCCTGTCTACCTCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220
QY 944 GCAGCGTGGGCGCCACCGAGCAGCAGAGGCTGTGGCGGCTGTGGCGGCTGTGGCGGCTGT 1003
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240
QY 1004 TGTCTGCATGCGCGCGCTGTGGCTTTCGGAACTGTGGCGGCTGTGGCTGTGGCTGTGGCTG 1063
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260
QY 1064 CACAGCTGTGCTGCCGATGCCCGGCTGTGGCGGCTGTGGCTGTGGCTGTGGCTGTGGCTG 1123
```

```
Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280
QY 1124 AGCTGGATGGCACTCATGACCTTCACCGCTGTTTACAGGATTTCGTGGCGAGGGCGTG 1183
Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheThrLeuPheValGlyGluGlyLeu 300
QY 1184 TACCAGGGCGTGGCCAGAGTGAGCCGGGACCGAGCCGGGACACACTATGATGAAGGC 1243
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320
QY 1244 GTTCGGATGGCACTGGCGCTTCCTCCAGTGGCCCATCTCCCTGGTCTCTCTCTG 1303
Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340
QY 1304 GTCATGACCGGTGTGTCAGCAGTTCGGCACGTCGACGCTCTATTGGCCAGTGGCA 1363
Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360
QY 1364 GCTTCCCTGGTGGCGGTGGCCACATGCCTGTCCACAGTGTGGCGGTGGTGAGCT 1423
Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380
QY 1424 TCAGCCCGCTCACCGGTTTCACCTCTCAGCCCTGCAGATCCTGCCCTACACATGGCC 1483
Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400
QY 1484 TCCTCTACACCGGGAGAGCAGGTGTCTCTGCCCAATACCGAGGGACACTGGAGT 1543
Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420
QY 1544 GCTAGCAGTCAGACAGCCTGATGACCATCTCTCCAGGCCCTAAGCCTGGAGCTGCC 1603
Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440
QY 1604 TTCCTTAATGGACACAGTGGTGTGGAGGAGTGGCTGTCTCCACCTCCACCGCGGCTC 1663
Db 441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 460
QY 1664 TGGGGGCTCTCCCTGTGATGTCTCCGTACGTGTGGTGGTGGGAGGAGCCACCGAGGCT 1723
Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480
QY 1724 AGGTGTGTCCGGCGCGGCACTCTCCGTGGACCTCGCCATCTGGATAGTGCCTCTCTG 1783
Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500
QY 1784 CTGTCCAGGTGCCCCATCCCTGTTTATGGCTCCATTGTCCAGTCCAGCCAGCTGTC 1843
Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520
QY 1844 ACTGCCCTATATGTGTCTGCCGAGCCTGGGTCTGGTGGCCATTACTTCTCTACACAG 1903
Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540
QY 1904 GTAGTATTTGACAGAGGACTTGGCCAAATACTCAGCG 1942
Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

## RESULT 13

```
US-09-895-793-947
; Sequence 947, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Reiter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
```

```
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.53AC2
; CURRENT APPLICATION NUMBER: US/09/895,793
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 947
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-947
```

## Alignment Scores:

```
Pred. No.: 3,26e-151 Length: 1079
Score: 2861.00 Matches: 553
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.58% Indels: 0
DB: 9 Gaps: 0
```

US-09-759-143-110 (1-3410) x US-09-895-793-947 (1-1079)

```
QY 284 ATGTTCCAGAGCGCTGTGGTGTAGCGCGCTGTGGCGACCGGAAAGCCAGCTTTGCTG 343
Db 527 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 546
QY 344 GTCACCTGCTTAACCTTTGGCTGTGGAGTGTGTTGGCGCAGGCATCACCTATGTCCG 403
Db 547 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyrValPro 566
QY 404 CCTCTGCTGTGGAGTGGGGTAGAGAGAGTTCATGACCATGTGCTGGCGATTGGT 463
Db 567 ProLeuLeuLeuGluValGlyValGluLysPheMetThrMetValLeuGlyIleGly 586
QY 464 CCAGTGTGGCGCTGGTGTGTCTCCGCTCTCTAGCTCAGCCAGTGCAGCTGCGGTGA 523
Db 587 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 606
QY 524 CGCTATGGCGCGCGCGCTTCATCTGGGCACTGTCTGGGCATCTCTGTCAGGCTC 583
Db 607 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 626
QY 584 TTTCTCATCCCAAGCGCGCTGTGTAGCAGGCTGTGTGTCGCGGATCCAGGCCCTG 643
Db 627 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 646
QY 644 GAGTGGCACTGCTCATCTGGGGCTGGGGCTGTGTGCTGTGTGTGTGTGTGTGTGTG 703
Db 647 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 666
QY 704 ACTCCACTGGAGGCGCTCTCTGACCTCTCTCGGACCGCGGACCGGACCACTGTCCGAGGCC 763
Db 667 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 686
QY 764 TACTCTCTATGCTTTCATCATCAGTCTTTGGGGCTGTGTGGGCTACCTCTCTGCTGCC 823
Db 687 TyrSerValTyrAlaPheMetIleSerLeuGlyCysLeuGlyTyrLeuLeuProAla 706
QY 824 ATTGACTGGGACACCACTGCGCTGGCGCTTACCTGGCGACCGGAGGAGAGTGCCTTTT 893
```

Db	707	IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe	726
Qy	884	GCCTGCTCACCTCATCTCTCACTACCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAG	943
Db	727	GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu	746
Qy	944	GCAGCGCTGGCCGCCACGAGCCAGAGAGGGGTGTGGCCGCCCTCTGTGTCGCCCCAC	1003
Db	747	AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis	766
Qy	1004	TGCTGTCCATGCGCGCGCTGTGCTTCCGGAACCTGGCGCCCTGCTTCCCGGCTG	1063
Db	767	CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu	786
Qy	1064	CACCACTGTGTGCGCATGCCCGCACCCCTGCGCGGCTTCTTGGTGGTGAAGTGTGC	1123
Db	787	HisGlnLeuCysCysArgMetProArgThrLeuArgLeuPheValAlaGluLeuCys	806
Qy	1124	AGCTGATGGCACTCATGACCTTCACGCTGTTTACAGGATTTCTGGCGAGGGGCTG	1183
Db	807	SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu	826
Qy	1184	TACAGGGGTGCCAGAGCTGAGCGCGGACCCAGGCGCGAGACACATATGATGAAGGC	1243
Db	827	TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly	846
Qy	1244	GTTCCGATGGCAGCGCTGGGCTGTCTGTCAGTGGCGCATCTCCCTGGTCTTCTCCTG	1303
Db	847	ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu	866
Qy	1304	GTCATGACCGGCTGTGTCAGCGATTGCGCACTCGAGCAGTCTATTGGCCAGTGTGGCA	1363
Db	867	ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla	886
Qy	1364	GCCTTCCCTGTGCTGCCGGTGCACATGCCCTGCCACAGTGGCGCTGGTGACAGCT	1423
Db	887	AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValAlaThrAla	906
Qy	1424	TCAGCGCCCTCACCGGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC	1483
Db	907	SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla	926
Qy	1484	TCCTCTACACCGGAGAGCAGGTGTTCTTGCCCAATATCCGAGGGGACACTGGAGGT	1543
Db	927	SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly	946
Qy	1544	GCTAGCAGTCAGACAGCCTGATGACAGCTTCTGCCAGGCCCTAAGCTGGAGCTCCC	1603
Db	947	AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro	966
Qy	1604	TTCCCTAATGGACACGTGGTGTGGAGCAGTGGCTGCTCCCACCTCCACCGCGTCC	1663
Db	967	PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeu	986
Qy	1664	TGCGGGCCCTGCCTGTGATGCTCCGTACGTGTGGTGGGTGAGCCACCCAGGCGC	1723
Db	987	CysGlyAlaSerAlaCysAspValSerValArgValValGlyGluProThrGluAla	1006
Qy	1724	AGGCTGTGTCGGGCGCGGCATCTGCTGGACCTCGCCATCTGGATAGTCCCTTCCTG	1783
Db	1007	ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu	1026
Qy	1784	CTGTCCAGGTGGCCCCATCCCTGTGTTATGGGCTCCATGTCCAGCTCAGCAGTCTGTC	1843
Db	1027	LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal	1046
Qy	1844	ACTGCCATATGTGTCTGCGCAGGCGCTGGGTCTGGTCCGCCATTTTACTTTGCTACACAG	1903
Db	1047	ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln	1066
Qy	1904	GTAGTATTGTGAAGAGCGACTTGGCCAAATACTCAGCG	1942

```
Db 727 GlyLeuLeuThrLeuPheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 746
QY 944 GCAGCGCTGGCCGACAGCCAGCAGAGGCTGTGGCCGCTCTTGTGCTGCCCCAC 1003
Db 747 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 766
QY 1004 TGCTGTCCATCCGGCCGCTTGGCTTCCGGAACCTGGCGCCCTGCTTCCCGGCTG 1063
Db 767 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 786
QY 1064 CACCAGCTGTCTCCGCGATGCCCGCCGACCTGCGCGGCTCTTCTGGCTGAGCTGTG 1123
Db 787 HisGlnLeuCysCysArgMetProArgThrLeuArgGluLeuPheValAlaGluLeuCys 806
QY 1124 AGCTGGATGGCACTCATGCTTCCAGCTGTCTTACACGGATTTCGTGGCGAGGGCTG 1183
Db 807 SerTrpMetAlaLeuMetThrPheThrLeuPheThrAspPheValGlyGluGlyLeu 836
QY 1184 TACCAGGCGTGGCCAGAGCTGAGCGGCGCAGCGAGCGCCGAGACACTATGATGAGGC 1243
Db 827 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTrpAspGluGly 846
QY 1244 GTTCGGATGGCAGCCTGGGCTGTCTCTGCACTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
Db 847 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 866
QY 1304 GTCATGACCGCTGGTGCAGGATTCGGCACTCGGACGCTATTTGGCCAGTGGGA 1363
Db 867 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTrpLeuAlaSerValAla 886
QY 1364 GCTTTCCTCTGGCTGGCGTGCACATGCTCTCCACAGTGTGGCGTGGTGACAGCT 1423
Db 887 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 906
QY 1424 TCAGCCGCTCACCAGGCTTACCTCTCAGCCCTGCAGATCCTGCCTACACTGGCC 1483
Db 907 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 926
QY 1484 TCCTCTTACCACCGGAGAGAGGTTCCTGCCCAATACCGAGGCGACACTGGAGT 1543
Db 927 SerLeuThrHisArgGluGlyGlnValPheLeuProLysTrpArgGlyAspThrGlyGly 946
QY 1544 GCTAGCAGTGGAGACAGCCTGTAGTACAGCTTCTGCCAGGCGCTAAGCCTGGAGTCCC 1603
Db 947 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 966
QY 1604 TTCCTAATGGACACGTGGTGTGAGCGAGTGTGAGCGCTCTCCAGCTCCACCGCGCTC 1663
Db 967 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 986
QY 1664 TGGCGGCGCTCTCCCTGTGATGCTCTCCGTACGTGTGTGGTGGTGGTGGTGGTGGTGG 1723
Db 987 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 1006
QY 1724 AGGCTGGTTCGGCGCGGCGATCTGCTGGACCTCGCCATCTGGATAGTCCCTCTCTG 1783
Db 1007 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 1026
QY 1784 CTGTCCAGGTGGCCCATCTCTGTTATGGGCTCCATGCTCCAGCTCAGCCAGTCTGTC 1843
Db 1027 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 1046
QY 1844 ACTGCTATATGCTGTCTCCGAGCGCTGTGGTCTGTGGCTGCTTACTTTGCTACACAG 1903
Db 1047 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 1066
QY 1904 GTAGTATTTGACAGAGCGACTTGGCCAAATACTCAGCG 1942
Db 1067 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 1079
RESULT 15
US-09-895-793-974
; Sequence 974, Application US/09895793
```

```
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Jennifer L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF PROSTATE CANCER
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 974
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-974

Alignment Scores:
Pred. No.: 1,35e-86 Length: 359
Score: 1696.00 Matches: 320
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.43% Indels: 0
DB: 9 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-895-793-974 (1-359)

QY 284 ATGTCCAGAGGCTGTGGTGTGAGCGCTGTGGCGGCAAGCCAGCCAGCTTTGCTG 343
Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20
QY 344 GTCAACCTGTCTAACCTTTTGGCTGTGGCTGTGGTGTGGTGTGGTGTGGTGTGG 403
Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyrValPro 40
QY 404 CCTCTGTCTGTGGAGTGGGGGTAGAGAGAAGTTTCATGACCATGTTGGTGGGCA 463
Db 41 ProLeuLeuGluValGlyValGlyValGlyValGlyValGlyValGlyValGly 60
QY 464 CCAGTGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 523
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArg 80
QY 524 CGCTATGCGCGCGCGCGCGCTTTCATCTGGGCACTGTCTTGGGCACTCTGCTGAG 583
Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100
QY 584 TTCTCTCATCCAGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 643
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgPro 120
QY 644 GAGCTGCACTGTCTATCTCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 703
```



Db 121 GluLeuAlaLeuLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
QY 704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGACCACTGCTCCAGGCC 763  
Db 141 ThrProLeuGluAlaLeuLeuSerAspPheArgAspProAspHisCysArgGlnAla 160  
QY 764 TACTCTGTATGCTCTCATGATCAGTCTTGGGGGCTGCTGGGCTACCTCCCTGCTGCC 823  
Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
QY 824 ATTGACTGGGACACAGTGCCTGCGCCCTACCTGGGACCCAGGAGAGTGCCTCTTT 883  
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe 200  
QY 884 GGCCTGTCCCTCATCTTCCCTCACCTGCTAGCAGCAGCAGTGTGGTGGTGGAG 943  
Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
QY 944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAGGGCTGTGGGCCCTCTTGTGCGCCCCAC 1003  
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
QY 1004 TGCTGTCCATGCGGGCCGCTTGGCTTTCGGGAACCTGGGGCCCTGCTTCCCGGCTG 1063  
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
QY 1064 CACGAGTGTGCTGCGCGCATGCCCGCACCTGCGCGCTCTCTGCTGGCTGAGCTGTC 1123  
Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
QY 1124 AGCTGGATGGCACTCATGACCTTTCACGCTGTTTACAGGATTTCTGGGCGAGGGCTG 1183  
Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
QY 1184 TACAGGCGTGGCCAGCTGACCGGCGGACCGGCGGAGCAGCAGTATGATGAGGC 1243  
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

## RESULT 16

US-09-822-827-974  
; Sequence 974, Application US/09822827  
; Patent No. US20020081680A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C1  
; CURRENT APPLICATION NUMBER: US/09/822,827  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 974  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-822-827-974

Alignment Scores:  
Pred. No.: 1,35e-86 Length: 359  
Score: 1696.00 Matches: 320  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 26.43% Indels: 0  
DB: 10 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-822-827-974 (1-359)

QY 284 ATGGTCCAGAGGCTGTGGTGCAGCGCTGCTGGGACCGGAAAGCCAGCTCTTCTG 343  
Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
QY 344 GTCAACCTGCTAACCTTTGGCTGGAGGTGTGTTGGCGGAGGCATCAGCTATGTCGG 403  
|||||

Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTrpValPro 40  
QY 404 CCTCTGCTCTGGAAGTGGGGTAGAGGAGAGTTCATGACCATGCTGCTGGGCATTTGGT 463  
Db 41 ProLeuLeuGluValGlyValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
QY 464 CCAGTCTGGGCTGTGCTGTGCTCCGCTCCTAGGCTCAGCCAGTACACACAGTGGCGTGA 523  
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
QY 524 CGCTATGGCGCGCGCGCTTCATCTGGGCACTGTCTCTGGGCACTCTCTGCTGAGCCTC 583  
Db 81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
QY 584 TTCTCATCCCAAGGCCCGCTGGTAGCAGGCTGCTGCTGGGATCCCGAGGCCCTG 643  
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
QY 644 GAGCTGGCACTGCTCATCTGCTGGGCTGCTGGACTTCTGTGGCAGGCTGTGCTTC 703  
Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
QY 704 ACTCCACTGGAGGCCCTGCTCTGACCTTTCGGGACCCGAGCAGTGTGCTCCAGGCC 763  
Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
QY 764 TACTCTGTATGCTCTCATGATCAGTCTTGGGGCTGCTGGGCTACCTCTCTGCTGCT 823  
Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
QY 824 ATTGACTGGGACACCACTGCTGCTGGCGCTTACCTGGGACCCAGGAGAGTGCCTCTTT 883  
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe 200  
QY 884 GGCTCTCACCTCTATCTTCTACCTGCTGCTAGCAGCAGCAGTGTGCTGCTGCTGAGGAG 943  
Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
QY 944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAGGGCTGTGGGCCCTCTCTGCTGCGCCAC 1003  
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
QY 1004 TGCTGTCCATGCGGGCCGCTTGGCTTTCGGGAACCTGGGGCCCTGCTTCCCGGCTG 1063  
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
QY 1064 CACGAGTGTGCTGCGCGCATGCCCGCACCTGCGCGCTCTCTGCTGGCTGAGCTGTC 1123  
Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
QY 1124 AGCTGGATGGCACTCATGACCTTTCACGCTGTTTACAGGATTTCTGGGCGAGGGCTG 1183  
Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
QY 1184 TACAGGCGTGGCCAGAGCTGAGCGGCGGACCGGCGGAGCAGCAGTATGATGAGGC 1243  
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

## RESULT 17

US-10-012-896-708  
; Sequence 708, Application US/10012896  
; Publication No. US20020183251A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.



Pred. No.: 2,57e-71 Length: 371  
Score: 1420.50 Matches: 275  
Percent Similarity: 88.58% Conservatives: 12  
Best Local Similarity: 84.88% Mismatches: 23  
Query Match: 22.13% Indels: 14  
DB: 9 Gaps: 1

US-09-759-143-110 (1-3410) x US-09-895-793-708 (1-371)

```
QY 347 AACCTGCTAACCTTTGGCTGGAGGTGTTTGGCGGAGGATCACCTATGTCGGCCT 406
Db 9 SerLeuValProLeuProLeuAlaLeuTyLeuSerGlnProLeuThrHisThrSer 28
QY 407 CTGCTGCTGGAAGTGGGGTAGAGGAGAGTTCATGACCATGCTGCTGGCATTGTCCTCA 466
Db 29 LeuLeu-----AlaGlylleGlyPro 35
QY 467 GTGCTGGGCTGTGCTGTCCGCTCCTAGGCTCAGCCAGTACACCTGGGTGGAGCC 526
Db 36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55
QY 527 TATGGCGCGCGCGCGCTTCATCTGGGCACTGCTGCTGGGCACTGCTGCTGACCTCTTT 586
Db 56 TyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyLeuLeuSerLeuPhe 75
QY 587 CTCATCCAAAGGCGCGCTGCTAGCAGGCTGCTGCTGGCGGATCCCGGCGGCTGGAG 646
Db 76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95
QY 647 CTGGCACTGCTATCTGCTGGCGCTGGGCTGCTGACTTCTGTGGCCAGGTGCTTCACT 706
Db 96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115
QY 707 CCAGTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766
Db 116 ProLeuGluAlaLeuLeuSerAspPheArgAspProAspHisCysArgGlnAlaTrp 135
QY 767 TCTGCTATGCTTCATGATCAGTCTTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 826
Db 136 SerValTyAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyLeuLeuProAlaIle 155
QY 827 GACTGGGACACAGTGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886
Db 156 AspTrpAspThrSerAlaLeuAlaProTyLeuGlyThrGlnGluCysLeuPheGly 175
QY 887 CTGCTACCTCTCATCTTCCCTCACCTGGTACGACCCACTGCTGCTGCTGCTGCTGCTGCT 946
Db 176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAla 195
QY 947 GCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1006
Db 196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215
QY 1007 TGTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1066
Db 216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235
QY 1067 CAGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1126
Db 236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255
QY 1127 TGGATGGCACTCATGACCTTCACGCTGTTTACAGGATTTGCTGGCGGAGGCGGCTGAC 1186
Db 256 TrpMetAlaLeuMetThrPheThrLeuPheTyThrAspPheValGlyGluGlyLeuTy 275
QY 1187 CAGGCGCTGCCAGAGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1246
Db 276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295
QY 1247 CGGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1306
Db 296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315
```

QY 1307 ATGACCGCG 1316  
Db 315 aValSerGly 318

## RESULT 19

US-09-895-814-708  
; Sequence 708, Application US/09895814  
; Publication No. US20020193296A1

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriek  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C26  
; CURRENT APPLICATION NUMBER: US/09/895,814  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 990  
; SOFTWARE: FASTSEQ FOR Windows Version 3.0  
; SEQ ID NO 708  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-895-814-708

## Alignment Scores:

Pred. No.: 2,57e-71 Length: 371  
Score: 1420.50 Matches: 275  
Percent Similarity: 88.58% Conservatives: 12  
Best Local Similarity: 84.88% Mismatches: 23  
Query Match: 22.13% Indels: 14  
DB: 9 Gaps: 1

US-09-759-143-110 (1-3410) x US-09-895-814-708 (1-371)

```
QY 347 AACCTGCTAACCTTTGGCTGGAGGTGTTTGGCGGAGGATCACCTATGTCGGCCT 406
Db 9 SerLeuValProLeuProLeuAlaLeuTyLeuSerGlnProLeuThrHisThrSer 28
QY 407 CTGCTGCTGGAAGTGGGGTAGAGGAGAGTTCATGACCATGCTGCTGGCATTGTCCTCA 466
Db 29 LeuLeu-----AlaGlylleGlyPro 35
QY 467 GTGCTGGGCTGTGCTGTCCGCTCCTAGGCTCAGCCAGTACACCTGGGTGGAGCC 526
Db 36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55
QY 527 TATGGCGCGCGCGCGCTTCATCTGGGCACTGCTGCTGGGCACTGCTGCTGACCTCTTT 586
Db 56 TyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyLeuLeuSerLeuPhe 75
QY 587 CTCATCCAAAGGCGCGCTGCTAGCAGGCTGCTGCTGGCGGATCCCGGCGGCTGGAG 646
Db 76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95
```

```

QY 647 CTGGCAGCTCATCTCTGGGGCTGGGGCTGTGTGACCTCTGTGCGCCAGGTGTGCTTCACT 706
Db LeuAlaLeuLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115
QY 707 CCAGTGGAGGCGCTCTCTGTGACCTCTTCCGGGACCCGAGCAGCAGTCTGCCAGGCTAC 766
Db ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135
QY 767 TCTGTCTATGCTTCATGATCAGTCTTGGGGGCTGCTGGGCTACCTCTCTGCTGCGCAT 826
Db SerValTyrAlaPheMetIleSerLeuGlyCysLeuGlyTyrLeuLeuProAlaIle 155
QY 827 GACTGGGACACAGTGGCTGGGGCTGCTGGGACCCAGAGAGTGCCTCTTTGGC 886
Db AsprtrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPheGly 175
QY 887 CTGCTCACCTCATCTTCTCCACCTGCTAGCAGCAGCAGTGTGTGGTGGTGGAGGGA 946
Db LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAla 195
QY 947 GCGCTGGGGCCCGACGAGCAGAGGCTGTGCGGCGGCTGTGCTGTGCGGCGGCTGC 1006
Db AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerProHisCys 215
QY 1007 TGTCCATGCGGGCGGCTTGGCTTCCGGAACCTGGGGCGCTGCTTCCCGGCTGCAC 1066
Db CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235
QY 1067 CAGCTGTGCTGCGGCTGCGGCTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1126
Db GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255
QY 1127 TGGATGGCATCATGACCTTACGCTGTTTACACCGATTTGCTGGCGGAGGCGGTAC 1186
Db TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlyLeuTyr 275
QY 1187 CAGGGGCTGCGGCGGCTGAGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246
Db GlnGlyValProArgAlaGluProGlyThrGluAlaArgHisTyrAspGluGlyLys 295
QY 1247 CGGATGGGAGCTGGGGCTGCTTCCGAGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCT 1306
Db AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProGluThrThrLeuGlyAl 315
QY 1307 ATGGACGGC 1316
Db 315 aValSerGly 318

```

## RESULT 20

US-09-759-143-708

Sequence 708, Application US/09759143

Patent No. US200202248A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

APPLICANT: Jilang, Yuqi

APPLICANT: Henderson, Robert A.

APPLICANT: Kalos, Michael D.

APPLICANT: Fanger, Gary R.

APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick

APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Hepler, William

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

```

FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 708
LENGTH: 371
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-143-708

```

## Alignment Scores:

```

Pred. No.: 2,57e-71 Length: 371
Score: 1420.50 Matches: 275
Percent Similarity: 88.58% Conservative: 12
Best Local Similarity: 84.88% Mismatches: 23
Query Match: 22.13% Indels: 14
DB: 10 Gaps: 1

```

US-09-759-143-110 (1-3410) x US-09-759-143-708 (1-371)

```

QY 347 AACCTGCTAACCTTTGGCTGGAGGTGTGTTGGCCGAGCATCACCTATGTGCCGCT 406
Db 9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28
QY 407 CTGCTGCTGGAGAGTGGGGGTAGAGAGAAGTTTCATGACCATGCTGGGCTTGGTCCA 466
Db 29 LeuLeu-----AlaGlyIleGlyPro 35
QY 467 GTGCTGGCGCTGGTCTGTGTCTCCGCTCTAGGCTCAGGACCTAGCAGTGGGAGGC 526
Db 36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55
QY 527 TATGCGCCCGCCGCGCTTTCATCTGGGCACTGCTTGGGCACTCTGCTGGGCTGCTTTT 586
Db 56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75
QY 587 CTCATCCCAAGGGCGGCTGGCTAGCAGGCTGTGTGCCGATCCAGGCCCTGGAG 646
Db 76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95
QY 647 CTGGCAGCTCATCTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 706
Db 96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115
QY 707 CCATGGAGGCGCTGCTCTGTGACCTCTTCGGGACCCGAGCAGTGTGCGGCGGCTAC 766
Db 116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135
QY 767 TCTGTCTATGCTTTCATGATCAGTCTTGGGGCTGGCTGGGCTGGCTGGGCTGGCT 826
Db 136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155
QY 827 GACTGGGACACAGTGGCGCTGGCGGCTTACCTGGGCGGCGGCGGCGGCGGCTTGGC 886
Db 156 AsprtrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPheGly 175
QY 887 CTGCTCACCTCATCTTCTCCACCTGCTAGCAGCAGCAGTGTGTGGTGGTGGAGGGA 946
Db 176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAla 195
QY 947 GCGCTGGGGCCCGACGAGCAGAGGCTGTGCGGCGGCTGTGCGGCGGCTGTGCGGCG 1006
Db 196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215
QY 1007 TGTCCATGCGGGCGGCTTGGCTTCCGGAACCTGGGGCGGCTGCTGCTGCTGCTGCTG 1066
Db 216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235
QY 1067 CAGCTGTGCTGCGGCTGCGGCTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1126
Db 236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

```

```
QY 1127 TGGATGGCACTACGACCTGTTTACAGGATTCGTGGCGGAGGGCTGTAC 1186
Db 256 TrpMetAlaLeuMetThrPheThrLeuPheThrPheValGlyGluGlyLeuTyr 275
QY 1187 CAGGCGTGGCCAGAGCTGACCGGGCCAGGAGCCGCGGACACTATGATGAAGCGCTT 1246
Db 276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295
QY 1247 CGGATGGCAGCGTGGGCTGTTCTGCAGTGGCCATCTCCCTGGTCTTCTCTGTGTC 1306
Db 296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProGluThrThrLeuGlyAl 315
QY 1307 ATGGACCGGC 1316
Db 315 avaisergly 318

RESULT 22
US-09-780-669-708
; Sequence 708, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 708
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-780-669-708

Alignment Scores:
Pred. No.: 2,57e-71 Length: 371
Score: 1420.50 Matches: 275
Percent Similarity: 88.58% Conservative: 12
Best Local Similarity: 84.88% Mismatches: 23
Query Match: 22.13% Indels: 14
DB: 1 Gaps: 1

US-09-759-143-110 (1-3410) x US-09-780-669-708 (1-371)
QY 347 AACCTGCTACCTTTGGCTGGAGGTGTGTTGGCCGCGGACCATCACTATGTGGCGCT 406
Db 9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28
QY 407 CTGCTGTGGAAGTGGGGGTAGAGAGAGTTCATGACCATGTGTGCTGGCGCATGTGCTCA 466
Db 29 LeuLeu-----AlaGlyIleGlyPro 35
QY 467 GTGCTGGGCGCTGCTGTGTGTCCTCTAGGCTCAGCCAGTGAACACTGGCGGTGACGC 526
```

```
Db 36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisThrPargGlyArg 55
QY 527 TATGGCGCGCGCGCGCTTCTCTGGGACACTGTCTTGGGACACTCTCTGCTGAGCGCTCTTT 586
Db 56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75
QY 587 CTCATCCCAAGGCGCGCTGGCTAGCAGGCGCTGCTGTGCCGATCCAGGCGCGCTGGAG 646
Db 76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95
QY 647 CTGGCACTGCTCATCTGGCGCTGGGCTGCTGGACTTCTGTGCCAGGTGCTGCTCACT 706
Db 96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115
QY 707 CCAGTGGAGCGCTGCTCTCTGACCTTCTCCGGAGCCCGGACCACTGTGCCAGGCGCTAC 766
Db 116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135
QY 767 TCTGTCTATGCCCTTCATGATCAGTCTTGGGGCTGCTGCTGGCTACCTCTGCTGCTGC 826
Db 136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155
QY 827 GACTGGGACACCACTGCTGCTGGCGCTGCTGCTGGCGCGCTGCTGCTGCTGCTGCTG 886
Db 156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPheGly 175
QY 887 CTGCTCACCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 946
Db 176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAla 195
QY 947 GCGTGGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1006
Db 196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215
QY 1007 TGTCCATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1066
Db 216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235
QY 1067 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1126
Db 236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255
QY 1127 TGGATGGCACTACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1186
Db 256 TrpMetAlaLeuMetThrPheThrLeuPheThrAspPheValGlyGluGlyLeuTyr 275
QY 1187 CAGGCGTGGCGCGAGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1246
Db 276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295
QY 1247 CGGATGGCAGCTGGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1306
Db 296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProGluThrThrLeuGlyAl 315
QY 1307 ATGGACCGGC 1316
Db 315 avaisergly 318

RESULT 22
US-09-822-827-708
; Sequence 708, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 708
```

; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-822-827-708

Alignment Scores: 2.57e-71 Length: 371  
Pred. No.: 1420.50 Matches: 275  
Score: 86.58% Conservative: 12  
Percent Similarity: 84.88% Mismatches: 23  
Best Local Similarity: 22.13% Indels: 14  
Query Match: 10 Gaps: 1  
DB: 1

US-09-759-143-110 (1-3410) x US-09-822-827-708 (1-371)

```
QY 347 AACCTGCTAACCTTTGGGCTGAGGTGTTGGCCGAGCATCACATATGTGCGGCT 406
      :::::  |||  ::  :::::
Db 9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrSer 28
      :::::  |||  ::  :::::
QY 407 CTGCTGCTGGAAGTGGGGTAGAGGAGAAGTTCATGACCATGGTGGCGATTGGTCCA 466
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 29 LeuLeu-----AlaGlyIleGlyPro 35
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 467 GTGCTGGGCTGGTGTGCTCCGCTCCTAGGCTCAGCCAGTACGACCTGGCGTGGAGCG 526
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 527 TATGGCGCGCGCGCTTCTCATCTGGGCACTGTCTTGGGCATCTCTGCTGAGCTCTTT 586
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 56 TyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 587 CTCATCCCAAGGCGCGCTGCTAGCAGGCTGTGTGCCGGATCCAGCCGCCCTGGAG 646
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 647 CTGGCACTGCTCATCTGGGCGTGGGCTGTGTGACTTCTGTGCCAGGTGTGCTTCACT 706
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 96 LeuAlaLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 707 CCATGGAGGCGCTCTCTGACCTCTTCGGGACCGGACCACTGTGCCAGGCTAC 766
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 116 ProLeuGluAlaLeuLeuSerAspLeuPheArgProAspHisCysArgGlnAlaTyr 135
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 767 TCTGTATGCTTCATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCCATT 826
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 827 GACTGGGACACCACTGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPheGly 175
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 887 CTGCTCACCTTCATCTTCTCCTCACCTGCTAGCAGCCACACTGTGTGGTGGAGGCA 946
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAla 195
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 947 CGCTGGGCCCCACCGAGCCAGCAGAGGCTGTGGCCCCCTCTGTGCTGCCCTGCTGC 1006
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1007 TGTCATGTCGGGCGCGCTTGGCTTCCGGAACCTGGGCGCTGCTGCTGCTGCTGCT 1066
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1067 CAGCTGTGCTGCCGATCCCGCCGACCTCTGCGCGGCTCTTCTGCTGCTGCTGCTGCT 1126
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1127 TGGATGGCACTCATGACCTTCACGCTGTTTACAGGATTTCTGGGCGAGGGCTGTAC 1186
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlyLeuTyr 275
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1187 CAGGGCTGCTGCCAGCTGAGCGGCGCACCGAGCGCGGAGACACTATGATGAGCGCT 1246
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
```

```
Db 276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295
QY 1247 CGGATGGGACGCTGGGCTGTTCTCTGCAGTGGCCCATCTCCCTGGTCTTCTCTCTGGTC 1306
      :::::  |||||  |||  :::::  |||  :::::  |||  :::::  |||  :::::
Db 296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProGluThrThrLeuGlyAl 315
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1307 ATGGACGGC 1316
      :  :::::
Db 315 aValSerGly 318
```

## RESULT 23

US-10-012-896-852  
; Sequence 852, Application US/10012896  
; Publication No. US20020183251A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Meagher, Madeleine Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C27  
; CURRENT APPLICATION NUMBER: US/10/012,896  
; CURRENT FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 1011  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 852  
; LENGTH: 400  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-012-896-852

Alignment Scores: 2.25e-70 Length: 400  
Pred. No.: 1403.50 Matches: 267  
Score: 86.58% Conservative: 0  
Percent Similarity: 97.45% Mismatches: 2  
Best Local Similarity: 22.87% Indels: 5  
Query Match: 9 Gaps: 1  
DB: 1

US-09-759-143-110 (1-3410) x US-10-012-896-852 (1-400)

```
QY 371 GTGCTTTGGCGCAGGC-----ATCACCTATGTGCCCTCTGCTGCTG 415
      |||  |||||  |||
Db 127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProLeuLeu 146
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 416 GAAGTGGGGTAGAGAGAAGTTTCATGACCATGGTGGTGGCCAGTGGCTGGGC 475
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 147 GluValGlyValGluGlyPheMetThrMetValLeuGlyIleGlyProValLeu 166
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 476 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186
      |||||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
```



QY 536 CGCGGCGCTTCATCTGGGCACTGTCTTGGGCACTCTCTGAGCTCTTTCTCATCCCA 595  
Db 187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206  
QY 596 AGGCGCGGCTGGCTAGCAGGCTGCTGTGCCCGATCCAGGCGCTTGGAGCTGGCACTG 655  
Db 207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluAlaLeu 226  
QY 656 CTCATCCCTGGGCGGCTGCTGGCACTCTGTGGCCAGGCTGCTTCACCTCCACTGGAG 715  
Db 227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246  
QY 716 GCCTGTCTCTGACTCTTCCGGGACCGGACCACTGTCCCGAGGCTACTCTGTCTAT 775  
Db 247 AlaLeuLeuSerAspLeuPheArgProAspHisCysArgGlnAlaTyrSerValTyr 266  
QY 776 GCCTTCATGATCAGTCTTGGGCGGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTG 835  
Db 267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286  
QY 836 ACCAGTGCCTGGGCGGCTTCTGCGGACCGGACCACTGTCCCGAGGCTACTCTGTCTAT 895  
Db 287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPheGlyLeuLeuThr 306  
QY 896 CTCATCTCTCCTCAGCTGTAGCAGCAGCACTGTCTGTGGCTGAGGAGGAGCGCTGGC 955  
Db 307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAlaAlaLeuGly 326  
QY 956 CCCACCGAGCAGCAGGAGGCTGTGGGCGGCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 1015  
Db 327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerLeuSerProHisCysProCys 346  
QY 1016 CGGCGCGCTGTGGCTTCCGGAACCTGGGCGGCTGCTTCCCGGCTGACCACTGCTGC 1075  
Db 347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366  
QY 1076 TGCCGATGCCCGGACCGCTGCGGCGGCTCTTGTGGCTGAGCTGTGAGCTGGATGCCA 1135  
Db 367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386  
QY 1136 CTCATGACCTTCAGCTCTTTTACACGATTTCTGGGCGAG 1177  
Db 387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400

## RESULT 24

US-09-895-793-852

Sequence 852, Application US/09895793

Publication No. US20020192763A1

## GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

APPLICANT: Jiang, Yuqiu

APPLICANT: Kalos, Michael D.

APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darriek

APPLICANT: Li, Samuel X.

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Hepler, William T.

APPLICANT: Henderson, Robert A.

APPLICANT: Hural, John

APPLICANT: McNeill, Patricia D.

APPLICANT: Houghton, Raymond L.

APPLICANT: Vinals de Bassols, Carlota

APPLICANT: Foy, Teresa

APPLICANT: Fanger, Gary R.

TITLE OF INVENTION: CONDITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.534C2

CURRENT APPLICATION NUMBER: US/09/895,793

CURRENT FILING DATE: 2001-06-29

NUMBER OF SEQ ID NOS: 982

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 852

LENGTH: 400

TYPE: PRT

ORGANISM: Homo sapiens

US-09-895-793-852

## Alignment Scores:

Pred. No.:	2,256-70	Length:	400
Score:	1403.50	Matches:	267
Percent Similarity:	97.45%	Conservative:	0
Best Local Similarity:	97.45%	Mismatches:	2
Query Match:	21.87%	Indels:	5
DB:	9	Gaps:	1

US-09-759-143-110 (1-3410) x US-09-895-793-852 (1-400)

QY 371 GTGTGTTTGGCCGAGGC-----ATCACCTATGTGCGGCTCTGCTGCTG 415  
Db 127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProLeuLeuLeu 146  
QY 416 GAAGTGGGCTACAGAGAGTTCATGACCATGCTGGCACTTGGTTCAGTCTGGGCG 475  
Db 147 GluValGlyValGluGluPheMetThrMetValLeuGlyIleGlyProValLeuGly 166  
QY 476 CTGCTGTGTCTCCGCTCTCTAGGCTCAGCCAGTGCAGCTGCGGCTGAGGCTATGCGCC 535  
Db 167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186  
QY 536 CGCGCGGCTTCATCTGGGCACTGTCTTGGGCACTCTGCTGAGCTCTTTCTCATCCCA 595  
Db 187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206  
QY 596 AGGCGCGGCTGTAGCAGGCTGTGCTCCCGATCCAGGCGCTGAGCTGGGCACTG 655  
Db 207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226  
QY 656 CTCATCTGTGGGCTGGGCTGTGTGGACTTCTGTGGCCAGGTGTCTTCTACCTCCACTGGAG 715  
Db 227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246  
QY 716 GCCTGTCTCTGACTCTTCCCGGACCGGACCACTGTGCGGCGCTACTCTGCTCTAT 775  
Db 247 AlaLeuLeuSerAspLeuPheArgProAspHisCysArgGlnAlaTyrSerValTyr 266  
QY 776 GCCTTCATGATCAGTCTTGGGCGCTGCTGGGCTACTCTCTGCTGCTGCTGCTGCTGCTG 835  
Db 267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286  
QY 836 ACCAGTGCCTTGGCGGCTTACCTGCGGACCGGAGGAGTGCCTCTTTGCGCTGCTGCTGCTG 895  
Db 287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306  
QY 896 CTCATCTCTCTGACTGCTAGCAGCAGCACTGCTGCTGGTGGCTGAGGAGGCGGCTGGG 955  
Db 307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAlaLeuGly 326  
QY 956 CCCACCGAGCAGCAGGAGGCTGCTGGGCGGCTCTCTTGTGCGGCGGCTGCTGCTGCTGCTG 1015  
Db 327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysProCys 346  
QY 1016 CGGCGCGCTTGGCTTTCGGAACCTGGGCGGCTGCTTCCCGGCTGAGCTGCTGCTGCTGCTG 1075  
Db 347 ArgAlaArgLeuAlaPheArgAsnLeuLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366  
QY 1076 TGCCGATGCCCGGACCGCTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135  
Db 367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386



US-09-780-669-852  
; Sequence 852, Application US/09780669  
; Patent No. US20020051977A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.

QY 836 ACCAGTGCCTGGCCCTACCTGGGACCCAGGAGAGTGCCTTTGGCCCTGCTCACC 895  
 Db 287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPheGlyLeuLeuThr 306  
 QY 896 CTCATCTTCCTACCTCGGTAGCAGCACACTGCTGGTGGCTGAGGAGGAGCGCTGGGC 955  
 Db 307 LeuilePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAlaLeuGly 326  
 QY 956 CCCACGAGCAGCAGAGGGCTGTGGGCCCTCTTGTGCCCCCACTGCTGCTCCATGC 1015  
 Db 327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHiscysCysProCys 346  
 QY 1016 CGGGCCCGCTGGCTTCCGGAACCTGGCGCCCTGCTCCCGGCTGACACAGCTGTGC 1075  
 Db 347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHiscInLeuCys 366  
 QY 1076 TGCCGATGCCCGCACCTGCGCCGCTCTTCTGCTGGCTGAGCTGTGCAGCTGGATGGCA 1135  
 Db 367 CysArgMetProArgThrLeuArgLeuPheValAlaGluLeuLeuSerTrpMetAla 386  
 QY 1136 CTCATGACCTTCAGGCTGTTTACAGGATTCCTGGGCGAG 1177  
 Db 387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400

## RESULT 28.

US-09-822-827-852  
 ; Sequence 852, Application US/09822827  
 ; Patent No. US20020081680A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
 ; FILE REFERENCE: 210121.534C1  
 ; CURRENT APPLICATION NUMBER: US/09/822,827  
 ; NUMBER OF SEQ ID NOS: 982  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 852  
 ; LENGTH: 400  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-822-827-852

Alignment Scores:  
 Pred. No.: 2,25e-70 Length: 400  
 Score: 1403.50 Matches: 267  
 Percent Similarity: 97.45% Conservatives: 0  
 Best Local Similarity: 97.45% Mismatches: 2  
 Query Match: 21.87% Indels: 5  
 DB: 10 Gaps: 1

US-09-759-143-110 (1-3410) x US-09-822-827-852 (1-400)

QY 371 GTGTGTTTGGCCGACGCGC-----ATCACCATTGTCGCCCTCTGCTGCTG 415  
 Db 127 ValThrLeuAlaGluGlyProAlaGluPheLeuThrTyrValProLeuLeuLeu 146  
 QY 416 GAAGTGGGGGTAGAGAGAGTTTCATGACCATGGTGGCATGGTCCAGTCTGCTGGC 475  
 Db 147 GluValGlyValGluGlyPheMetThrMetValLeuGlyIleGlyProValLeuGly 166  
 QY 476 CTGGTGTGTGCTCCGCTCCTAGGCTCAGCCAGTACAGGCTGGCGTGGAGCTATGGCGC 535  
 Db 167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186  
 QY 536 CGCGCGCCCTTCATCTGGGCATCTGCTTGGGCATCTGCTGAGCCTCTTCTCATCCCA 595  
 Db 187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206  
 QY 596 AGGGCCGCTGCTAGCAGGCTGTGTGCGCGGATCCAGGCGCCCTGGAGCTGGCACTG 655  
 Db 207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

QY 656 CTCATCTGCGGCTGGGCTGCTGAGCTTCTGTGCCAGGTGCTGCTTCACTCCACTGGAG 715  
 Db 227 LeuileLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246  
 QY 716 GCCCTGCTCTGAGCTCTTCCGGACCCCGGACCACTGTCCGCCAGGCTACTTCTCTAT 775  
 Db 247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266  
 QY 776 GCCTTCATGATCAGTCTTGGGGCTGCTGGGCTACCTTCCTGCCCTGCCCTGACTGGGAC 835  
 Db 267 AlaPheMetIleSerLeuGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpasp 286  
 QY 836 ACCAGTGCCTGGCCCTTACCTGGGCACCCAGGAGAGTGCCTTTTGGCTGCTCACC 895  
 Db 287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPheGlyLeuLeuThr 306  
 QY 896 CTCATCTTCTCACCCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAGGAGCGCTGGGC 955  
 Db 307 LeuilePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaLeuGly 326  
 QY 956 CCCACGAGCAGCAGAGGGCTGTGGGCCCTCTTGTGCCCCCACTGCTGCTCCATGC 1015  
 Db 327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHiscysCysProCys 346  
 QY 1016 CGGGCCCGCTGGCTTCCGGAACCTGGCGCCCTGCTCCCGGCTGACACAGCTGTGC 1075  
 Db 347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHiscInLeuCys 366  
 QY 1076 TGCCGATGCCCGCACCTGCGCCGCTCTTCTGCTGGCTGAGCTGTGCAGCTGGATGGCA 1135  
 Db 367 CysArgMetProArgThrLeuArgLeuPheValAlaGluLeuLeuSerTrpMetAla 386  
 QY 1136 CTCATGACCTTCAGGCTGTTTACAGGATTCCTGGGCGAG 1177  
 Db 387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400

## RESULT 29

US-10-012-896-1011  
 ; Sequence 1011, Application US/10012896  
 ; Publication No. US20020183251A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yuqiu  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Retter, Marc W.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Darlick  
 ; APPLICANT: Li, Samuel X.  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Hepler, William T.  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Hural, John  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Houghton, Raymond L.  
 ; APPLICANT: Vinals de Bassols, Carlota  
 ; APPLICANT: Foy, Teresa  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Watanabe, Yoshihiro  
 ; APPLICANT: Meagher, Madeleine Joy  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
 ; FILE REFERENCE: 210121.427C27  
 ; CURRENT APPLICATION NUMBER: US/10/012,896  
 ; CURRENT FILING DATE: 2001-12-10  
 ; NUMBER OF SEQ ID NOS: 1011  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1011

FILE REFERENCE: 210121.427D3  
CURRENT APPLICATION NUMBER: US/10/010,940  
CURRENT FILING DATE: 2001-12-05  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 571  
LENGTH: 84  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-012-896-1011

## Alignment Scores:

Pred. No.: 4,66e-45 Length: 355  
Score: 947.00 Matches: 188  
Percent Similarity: 99.47% Conservative: 0  
Best Local Similarity: 99.47% Mismatches: 1  
Query Match: 14.76% Indels: 0  
DB: 9 Gaps: 0

US-09-759-143-110 (1-3410) x US-10-012-896-1011 (1-355)

QY 1307 ATGGACGGCTGGTGGACGGATTCGGGACTCGAGCAGTCTATTGGCCAGTGTGGCAGCT 1366  
DB 160 MetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAla 179  
QY 1367 TTCCTGTGGCTGCCGGTGCCACATGCTGCCACAGTGTGGCGGTGGTGCACAGCTTCA 1426  
DB 180 PheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValAlaThrAlaSer 199  
QY 1427 GCGGCCCTCACCGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCCTCC 1486  
DB 200 AlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSer 219  
QY 1487 CTTACACCGGAGAGCAGTGTCTGCTGCCCAATACCGAGGGGACACTGGAGTCT 1546  
DB 220 LeuTyrHisArgGlnLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAla 239  
QY 1547 AGCAGTGGAGACAGCTGTATGACAGCTTCTCCGCCAGGCCCTTAAGCCTGGAGCTCCCTC 1606  
DB 240 SerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPhe 259  
QY 1607 CCTAATGGACAGTGGGTCTGGAGCAGTGGCTGCTCCACCTCCACCGCCGCTGTC 1666  
DB 260 ProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeuCys 279  
QY 1667 GGGGCTCTGCTGTCTCTCTCCGTACGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1726  
DB 280 GlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAlaArg 299  
QY 1727 GTGGTTCGGGGCGGGGATCTGCTGGACCTGCTGGACCTGCTGGATGCTGGTGGTGGTGGT 1786  
DB 300 ValValProGlyArgGlyLeuCysLeuAspLeuAlaIleLeuAspSerAlaPheLeuLeu 319  
QY 1787 TCCAGGTGGCCCATCCCTGTTTATGGGCTCCATGTCAGCTCCAGCTCAGCCAGTGTCTACT 1846  
DB 320 SerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerValThr 339  
QY 1847 GCCTATATGGTGTCTCCCGCAGGCTG 1873  
DB 340 AlatyMetValSerAlaAlaLeu 348

## RESULT 30

US-10-010-940-571  
Sequence 571, Application US/10010940  
Publication No. US2003008062A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang Yuqi  
APPLICANT: Reed, Steven G.  
APPLICANT: Kalos, Michael  
APPLICANT: Fanger, Gary  
APPLICANT: Retter, Mark  
APPLICANT: Solk, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427D3  
CURRENT APPLICATION NUMBER: US/10/010,940  
CURRENT FILING DATE: 2001-12-05  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 571  
LENGTH: 84  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-010-940-571

Alignment Scores:  
Pred. No.: 1.37e-17 Length: 84  
Score: 452.00 Matches: 84  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.04% Indels: 0  
DB: 9 Gaps: 0

US-09-759-143-110 (1-3410) x US-10-010-940-571 (1-84)

QY 1493 CACCGGAGAGCAGGTGTTCTCCCAATACCGAGGGGACACTGGAGTGTAGCAGT 1552  
DB 1 HisArgGlnLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAlaSerSer 20  
QY 1553 GAGGACAGCCTGATGACAGCTTCTGCCAGGCCCTTAAGCCTGGAGCTCCCTTCCCTTAAT 1612  
DB 21 GluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPheProAsn 40  
QY 1613 GGACAGCTGGTGTGGAGGAGTGGCTGCTCCACCTCCACCGCGCTCTGGGGGGCC 1672  
DB 41 GlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeuCysGlyAla 60  
QY 1673 TCTGCTGTGATGCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1732  
DB 61 SerAlaCysAspValSerValArgValValValGlyGluProThrGluAlaArgValVal 80  
QY 1733 CCGGGCGGGGCG 1744  
DB 81 ProGlyArgGly 84

## RESULT 31

US-10-012-896-706  
Sequence 706, Application US/10012896  
Publication No. US20020183251A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Kalos, Michael D.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darriek  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals de Bassols, Carlota  
APPLICANT: Foy, Teresa  
APPLICANT: Fanger, Gary R.  
APPLICANT: Wantanabe, Yoshihiro  
APPLICANT: Meagher, Madeleine Joy  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C27

```
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-706

Alignment Scores:
Pred. No.: 3,74e-16 Length: 123
Score: 426.00 Matches: 87
Percent Similarity: 97.83% Conservatives: 3
Best Local Similarity: 94.57% Mismatches: 2
Query Match: 6.64% Indels: 0
DB: Gaps: 0

US-09-759-143-110 (1-3410) x US-10-012-896-706 (1-123)
QY 1250 ATGGGAGCGCTGGGCTGTTCTGCAGTGGCCCATCTCCCTGGTCTTCTCTGTGTCATG 1309
Db 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20
QY 1310 GACCGGCTGGTGCAGCGATTGCGCACTCGAGCAGTCTATTGGCCAGTGTGGCAGCTTTC 1369
Db 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40
QY 1370 CCTGTGGCTGGCGGTGCCACATGCTGTGCCACAGTGTGGCGTGGTGCAGCTTCAGCC 1429
Db 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAlaSerAla 60
QY 1430 GCCTCACCAGGTTACCTTCTCAGCCCTGCAGATCTCCCTACACACTGGCCCTCCCTC 1489
Db 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80
QY 1490 TACCACCGGAGAGCAGGTGTTCTGCCCAATAC 1525
Db 81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 32
US-09-895-793-706
; Sequence 706, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.53AC2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982

US-09-759-143-110 (1-3410) x US-09-895-793-706 (1-123)
QY 1250 ATGGGAGCGCTGGGCTGTTCTGCAGTGGCCCATCTCCCTGGTCTTCTCTGTGTCATG 1309
Db 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20
QY 1310 GACCGGCTGGTGCAGCGATTGCGCACTCGAGCAGTCTATTGGCCAGTGTGGCAGCTTTC 1369
Db 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40
QY 1370 CCTGTGGCTGGCGGTGCCACATGCTGTGCCACAGTGTGGCGTGGTGCAGCTTCAGCC 1429
Db 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAlaSerAla 60
QY 1430 GCCTCACCAGGTTACCTTCTCAGCCCTGCAGATCTCCCTACACACTGGCCCTCCCTC 1489
Db 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80
QY 1490 TACCACCGGAGAGCAGGTGTTCTGCCCAATAC 1525
Db 81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 33
US-09-895-814-706
; Sequence 706, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 706
; LENGTH: 123
```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-895-814-706

Alignment Scores: 3.74e-16 Length: 123  
Pred. No.: 426.00 Matches: 87  
Score: 97.83% Conservative: 3  
Best Local Similarity: 94.57% Mismatches: 2  
Query Match: 6.64% Indels: 0  
DB: 9 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-895-814-706 (1-123)

QY 1250 ATGGGAGCGCTGGGGCTGTTCTCGCAGTGGCCATCTCCCTGGTCTTCTCTGTCATG 1309  
DB 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20  
QY 1310 GACCGGTGGTCAGCGATTTCGCACTCGAGCAGTCTATTGGCCAGTGGCAGCTTTC 1369  
DB 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerAlaPhe 40  
QY 1370 CCTGTGCTGCCGTGCCATGCCCTGTCCACAGTGTGGCCGTGTGACAGCTTCAGCC 1429  
DB 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60  
QY 1430 GCCCTACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCCTCC 1489  
DB 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80  
QY 1490 TACCACCGGAGAGCAGGTGTTCTCCCTGCCCAATAC 1525  
DB 81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

## RESULT 34

US-09-759-143-706 ; Sequence 706, Application US/09759143  
Patent No. US200202248A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqui  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C23  
CURRENT APPLICATION NUMBER: US/09/759,143  
CURRENT FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 934  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 706  
LENGTH: 123  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-759-143-706

Alignment Scores: 3.74e-16 Length: 123  
Pred. No.: 426.00 Matches: 87  
Score: 97.83% Conservative: 3

Best Local Similarity: 94.57% Mismatches: 2  
Query Match: 6.64% Indels: 0  
DB: 10 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-759-143-706 (1-123)

QY 1250 ATGGGAGCGCTGGGGCTGTTCTCGCAGTGGCCATCTCCCTGGTCTTCTCTGTCATG 1309  
DB 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20  
QY 1310 GACCGGTGGTCAGCGATTTCGCACTCGAGCAGTCTATTGGCCAGTGGCAGCTTTC 1369  
DB 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40  
QY 1370 CCTGTGCTGCCGTGCCATGCCCTGTCCACAGTGTGGCCGTGTGACAGCTTCAGCC 1429  
DB 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60  
QY 1430 GCCCTACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCCTCC 1489  
DB 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80  
QY 1490 TACCACCGGAGAGCAGGTGTTCTCCCTGCCCAATAC 1525  
DB 81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

## RESULT 35

US-09-780-669-706 ; Sequence 706, Application US/09780669  
Patent No. US20020051977A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqui  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C24  
CURRENT APPLICATION NUMBER: US/09/780,669  
CURRENT FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 943  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 706  
LENGTH: 123  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-780-669-706

Alignment Scores: 3.74e-16 Length: 123  
Pred. No.: 426.00 Matches: 87  
Score: 97.83% Conservative: 3  
Best Local Similarity: 94.57% Mismatches: 2  
Query Match: 6.64% Indels: 0  
DB: 10 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-780-669-706 (1-123)



QY 1250 ATGGGAGCGTGGGGTGTCTCCAGTCAGCGCCCATCTCCCTGGTCTCTCTCTGGTCATG 1309  
Db 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20  
QY 1310 GACCGGTGTGGTGCAGGATTCGGACATCGAGCAGTCTATTGGCCAGTGTGGCAGCTTTC 1369  
Db 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40  
QY 1370 CCTGTGGCTGCCGTGCCACATGCTCTCCACAGTGTGGCCGTGGTGCAGCTTCAGCC 1429  
Db 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60  
QY 1430 GCCCTCACCGGTTACCTTCTCAGCCCTGCAGCTCTGCCCTACACACTGGCCTCCCTC 1489  
Db 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80  
QY 1490 TACCACCGGAGAGCAGGTGTCCTGCCCAATAC 1525  
Db 81 TyrHisArgGlnValLeuIleGlyGlnTrp 92

## RESULT 36

US-09-822-827-706

; Sequence 706, Application US/09822827

; Patent No. US20020081680A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.534C1

; CURRENT APPLICATION NUMBER: US/09/822,827

; CURRENT FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 706

; LENGTH: 123

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-822-827-706

Alignment Scores:  
Pred. No.: 3-74e-16 Length: 123  
Score: 426.00 Matches: 87  
Percent Similarity: 97.83% Conservative: 3  
Best Local Similarity: 94.57% Mismatches: 2  
Query Match: 6.64% Indels: 0  
DB: 10 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-822-827-706 (1-123)

QY 1250 ATGGGAGCGTGGGGTGTCTCCAGTCAGCGCCCATCTCCCTGGTCTCTCTCTGGTCATG 1309  
Db 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20  
QY 1310 GACCGGTGTGGTGCAGGATTCGGACATCGAGCAGTCTATTGGCCAGTGTGGCAGCTTTC 1369  
Db 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40  
QY 1370 CCTGTGGCTGCCGTGCCACATGCTCTCCACAGTGTGGCCGTGGTGCAGCTTCAGCC 1429  
Db 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60  
QY 1430 GCCCTCACCGGTTACCTTCTCAGCCCTGCAGCTCTGCCCTACACACTGGCCTCCCTC 1489  
Db 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80  
QY 1490 TACCACCGGAGAGCAGGTGTCCTGCCCAATAC 1525  
Db 81 TyrHisArgGlnValLeuIleGlyGlnTrp 92

## RESULT 37

US-10-096-986-103

; Sequence 103, Application US/10096986

Publication No. US20030083464A1  
GENERAL INFORMATION:  
APPLICANT: Ferrari, Franco A.  
Richardson, Charles  
Chambers, James  
Cauley, Stuart  
Pollock, Thomas J.  
Cappello, Joseph  
Crissman, John W.  
TITLE OF INVENTION: No. US20030083464A1 Peptides Comprising Repetitive  
Units of Amino Acids and DNA Sequences Encoding the Same  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr Hombach Test Albritton & Herbert LLP  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/096,986  
FILING DATE: 12-Mar-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/444,791  
FILING DATE: 22-No. US20030083464A1-1999  
APPLICATION NUMBER: US 08/482,085  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/175,155  
FILING DATE: 29-DEC-1993  
APPLICATION NUMBER: US 08/053,049  
FILING DATE: 22-APR-1993  
APPLICATION NUMBER: US 07/114,618  
FILING DATE: 29-OCT-1987  
APPLICATION NUMBER: US 06/927,258  
FILING DATE: 04-NOV-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-55186-11/RFT/BTC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 837 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
US-10-096-986-103  
Alignment Scores:  
Pred. No.: 2.54e-12 Length: 837  
Score: 356.50 Matches: 263  
Percent Similarity: 35.16% Conservative: 32  
Best Local Similarity: 31.35% Mismatches: 327  
Query Match: 5.73% Indels: 218  
DB: 9 Gaps: 57  
US-09-759-143-110 (1-3410) x US-10-096-986-103 (1-837)  
QY 2347 GGGAAACACAGTGACTGAGTTTATTCAGCTCCCAAAACCCCTTCTAGGTGTGTCTCAA 2288  
Db 46 GlySerProGly-----AlaProGlyThrPro-----GlyProGlnGly 58  
QY 2287 CTAGGAGGCTAGCTGTGTTAACCTCAGCCCTGGGTGTAATCCACCTGCAGAGTCCCGCATTC 2228





524	--GlyGluAlaGlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGly	542
1031	AAGCCAAACGGCGGCCGCATGGACAGCAGTGGGGGCAACAAGAGGGGGCCACAGCCCTT	972
543	SerProGlySerProGlyProAspGly	---
971	CTGCTGGCTGGTGGGCGCCAGG-	---
557	-----GlyProAlaGlyGlnAspGlyArgProGlyProProGlyProPro	571
929	-----GCAGTGTGGCTG	918
572	GlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyProLysGlyAlaAla	---589
917	CTAGCCAGGTGGAAGATCAGCGTGAGCAGGCCCAANGAGGCAC	---CCCTCTGGG
590	-----GlyGluProGlyLysAlaGlyGluArgGlyValProGlyProProGly	605
863	TGCCAGGT	---AGG
606	AlaValGlyProAlaGlyLysAspGlyGluAlaGlyAlaGlnGlyProProGlyProAla	625
851	GGGCCAGGGCACTGGTCTCCCAGTCATATGCCAGGCAGGA	---813
626	GlyProAlaGlyGluArgGlyGluGlnGlyProAlaGlySerProGlyPheGlnGlyLeu	645
812	---GSTATGCCAGCACGCCCCCAAGCATGATCATGAAGCATAGCAGTAGTAGGCTGGC	756
646	ProGlyProAlaGlyProProGlyGluAlaGlyLysProGlyGluGlnGlyValProGly	665
755	GACAGTGT	---CGGGTCCC
666	AspLeuGlyAlaProGlyProSerGlyAlaArgGlyGluArg	---GlyPheProGlyGlu
704	TGAACCACACTGCCACAGAAGTCCAGCAGCCCCAGCGCCAGGATGACAGTGCCA	---648
685	ArgGlyValGlnGlyProProGlyProAlaGlyProAlaGlnGlyAlaAsnGlyAlaProGly	704
647	-----GCTCCAGGGGCGCTGGATCCGGGCACACAGCC	615
705	AsnAspGlyAlaLysGlyAspAlaGlyAlaProGlyAla	---ProGlySerGlnGly
614	CTGCTACCCAGCGGCCCTTGGGTATGAAGAAGGCTCAGCAGGATGCCCAAGGACAGTG	555
723	AlaProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuProGly	---740
554	CCAGATGAAGGGCGCGCGGCCATAGCTCCAGCCAGTGGTCACTGGCTGAGCCCTA	495
741	ProLysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly	---SerPro
494	GGAGCG	---GNACACAGA
758	GlyLysAspGlyValArgGlyLeuThrGlyProIleGlyProProGlyProAla	---775
461	CAATGCCCA	---453
776	GlyAlaProGlyAspLysGlyLysSerGlyProSerGlyProAlaGlyProThrGlyAla	795
452	-----GCACCATGTGATGAACCTTCCTCTACCCCTTCCAGCAGCAGAGCGGCA	399
796	ArgGlyAlaProGlyAspArgGlyGluProGlyProProGlyProAlaGlyPheAlaGly	815
398	CATAGTGATGCTCGGCCCAACACACTCCAGC	---CAAAGTTAGCA
816	-----ProProGlyAlaAspGlyGlnProGlyAlaLys	826
350	GTTGTACCAAGAGCTGGGCTTTCGGTGCCAGCAGCGCGCTCACCCAGCCTCT	291
827	GlyGluProGlyAspAlaGlyAlaLysGlyAspAla	---GlyProProGlyProAla
290	GGACCATAGTGGGCCA	---GGCGGGTAGGCTCAGGGGGCGCTCAGGCAC
845	GlyPro-AlaGlyProGlyProIleGlyAsnValGlyAlaProGlyLysGlyAl	864

Qy	242	TCCAGAACTGCTTCGTCTCGCTCTGCTCT-----CCAGAACTGGCGCCTCTCTCTCTCTTG	189
Db	864	a-----ArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAl	879
Qy	188	CTGCCGCACAACGCTAGGAATCAGCCAGCGCCCATTTTGCAGCCCTTTGGTGCCCG	129
Db	879	aAlaGlyArgValGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProGl	899
Qy	128	TCCAGCTTCTCAAGCCCATGCTCAACACTGCTGCTGTGG-----GCACCTTCAGTGGGG	75
Db	899	yProAlaGlyLysglu-----GlyGlyLysglyProArg-GlyG	912
Qy	74	ACAGCTCTCATCACTCAGATCCTGGCCGA-----G	45
Db	912	luThrGlyProAlaGlyArgProGlyGluValGlyProProGlyProProGlyProProAlaG	932
Qy	44	GCGCGCGGCTGTCAACCGGAGCC	22
Db	932	lyGluLysGlySerProGlyAla	939
RESULT 39			
US-10-060-036-159			
; Sequence 159, Application US/10060036			
; Publication No. US2003007314A1			
; GENERAL INFORMATION:			
; APPLICANT: Benson, Darin R.			
; APPLICANT: Kalos, Michael D.			
; APPLICANT: Lodes, Michael J.			
; APPLICANT: Persing, David H.			
; APPLICANT: Hepler, William T.			
; APPLICANT: Jiang, Yuqiu			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
; FILE REFERENCE: 210121.566			
; CURRENT APPLICATION NUMBER: US/10/060.036			
; CURRENT FILING DATE: 2002-01-30			
; NUMBER OF SEQ ID NOS: 4560			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 159			
; LENGTH: 1464			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-060-036-159			
Alignment Scores:			
Pred. No.: 1.4e-11 Length: 1464			
Score: 343.00 Matches: 266			
Percent Similarity: 32.32% Conservative: 42			
Best Local Similarity: 27.91% Mismatches: 331			
Query Match: 5.51% Indels: 315			
DB: 9 Gaps: 55			
US-09-759-143-110 (1-3410) x US-10-060-036-159 (1-1464)			
Qy	2353	AGAGATGGGAACACAGGTGACTGAGTTTTATTTCAGTCCCAAAAACCCT-----TCTCTA	23000
Db	131	ArgAspGlyIleProGlyClnProGlyLeuProGlyProProGlyProPro 150	
Qy	2299	GGTGTCTCTCACTAGGAGGTAGCTGTTA---ACCCTGAGCTGGGTAAATCCACCTGCA	2243
Db	151	GlyProProGlyLeuGlyGlyAsnPheAlaProGlnLeuSerTyrrGlyTyrraspGluLys	170
Qy	2242	GAGTCCCCCATTCAGTCATCGAGCCCTTCTGGCCCTCCCTGTATAAGTCCAGACTCAA	218
Db	171	SerThrrGlyGlyIleSerValProGlyPrometGly-----	182
Qy	2182	ACCCCCTTGAAGGCCTCCAGTCAGCACGCCCTTAGAGACTGGGGGAGAGAGAGGACG	2123
Db	183	-----ProSerGlyPro-----ArgGlyLeu	189
Qy	2122	CCCCAGCCCCAGCTGTGCACTAGCACCTCAGCAGCACAGGGTGGCAGCAGAGCCCA	2063

Db 190 ProGlyProProGlyAlaPro---GlyProGlnGlyPheGlnGlyProProGlyGluPro 208  
QY 2062 CATTACTTTGGCAGCAACAGAACTGGCGCCAGCCCGCCGATGGGCTAACAGG 2003  
Db 209 GlyCluProGlyAlaSerGlyProMetGlyProAArgGlyProProGlyProGlyLys 228  
QY 2002 AGCGGGAGCTGGGA-----CCAGTGAGCGACCCCTCCA--- 1967  
Db 229 AsnGlyAspGlyGluAlaGlyLysProGlyAArgProGlyGluAArgGlyProGly 248  
QY 1966 -----CCCCAATGCTGGAGTTTCTACGCTGAGTATTGGC 1928  
Db 249 ProGlnGlyAlaArgGlyLeuProGlyThrAlaGly-----LeuProGly 263  
QY 1927 CAAGTCGCTCTTGTCAATATACCTGTGTAGCAAAAGTAAATGGCAGCAGACCCAGGCC 1868  
Db 264 MetLysGlyHisArgGlyPheSerGlyLeuAspGlyAlaLysGlyAspAlaGlyProAla 283  
QY 1867 TGGCGCAGACCATATAGCGAGTGCAGACTGGCTGGACATGGAGCCCATAAA 1808  
Db 284 GlyProLysGlyGluProGlySerProGly-----GluAsnGlyAlaProGly 299  
QY 1807 CAGGGATGGGCA-----CCTGGCAGCAGAGGACACTATCCAGATGGCGAG 1757  
Db 300 Gln--MetGlyProArgGlyLeuProGlyGluAArgGlyArg-----ProGlyAlaProG 317  
QY 1756 GTCCAGGCAGATGCCCGCCCGGAA-----CCACCCTGGCCTC 1718  
Db 317 LyProAlaGlyAlaAArgGlyAsnAspGlyAlaThrGlyAlaAlaGlyProProGlyProt 337  
QY 1717 GGTGGGCTCACCCACCACACACA-----CGTAGCGAGACATCAC 1680  
Db 337 hrGlyProAlaGlyProProGlyPheProGlyAlaValGlyAlaLysGlyGluAlaGlyP 357  
QY 1679 AG-----GCAGAGCCCGCAGAGCGCGGTGGAGTGGGAGCGGCCAC 1635  
Db 357 roGlnGlyProArgGlySerGluGlyProGlnGlyValAArgGly-----GluProG 374  
QY 1634 TGCTTCCAGCACCCAGCTGTCCATTAGG---AAGGGAGCTCCAGGCTTAGG---C 1584  
Db 374 lyProProGlyProAlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGlnP 394  
QY 1583 CTGGCAGGAAGTGTCTACAGCTGTCTCTCCTACCTAGCACCTCCAGTGTCCCTCGGT 1524  
Db 394 roGlyAlaLys-GlyAlaAsnGlyAlaPro-----Gly 404  
QY 1523 ATTTGGGAGGAAACACTGTCTTCTCCCGTGTGTAGAGGAGGCCAGTGTGTAGGCGAGGA 1464  
Db 405 ile---AlaGlyAlaProGlyPheProGlyAlaAArgGly-----ProSerGly 419  
QY 1463 TCTGCGAGGCTGAGAAGTGAACCGGTGAGGGCGGTGAAGCTGTCCACACGCGCCACAC 1404  
Db 420 Pro-GlnGlyProGlyProGlyProProGlyProGlyGlyAsnSerGlyGluProGly---Al 438  
QY 1403 TGTGGACAGCATGTGGCACCGGCGCAGCCAGGAAAGCTGCCACACTGGCCAAATAGA 1344  
Db 438 aProGlySerLysGlyAspThrGly-AlaLysGlyGluProGlyPro----- 453  
QY 1343 CTGCTCGAGTCCGNAATCGTGCACCGCCCGGTCCATGACACAG-----AGAAGACCA 1290  
Db 454 -----Val-GlyValGlnGlyProProGlyProAlaGlyGluGlyLysArgGly 470  
QY 1289 GGGAGATGGCGACTGCAGGAACAGCC---CCAGGCTGCCATCCGAAAGC----- 1242  
Db 471 AlaAArgGlyGluProGlyProThrGlyLeuProGlyProProGlyGluAArgGlyPro 490  
QY 1241 -----CTTCATCATAGTGTCTCCGGGCTCGGTCCCGGCTCAGCTC 1200  
Db 491 GlySerArgGlyPheProGlyAlaAspGlyAlaGlyProLysGlyProAlaGlyGlu 510  
QY 1199 TGGCAGCCCTGTACAGCCCTCGCCCAAGAAATCCGTGTAAACAGCGTGAAGTCA 1140  
Db 511 ArgGlySerProGlyProAlaGlyProLysGlySerPro----- 523

QY 1139 TGAGTGCCATCCAGCTGCACAGCTCAGCCACCAAGAGCGCGCGAGGTCTGGGGCATGC 1080  
Db 523 ----- 523  
QY 1079 GGCAGCAGCTGCTGTGACGGGGAAGAGCGCGCCAGCT-----TCCGGA 1032  
Db 524 ---GlyGluAlaGlyAArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGly 542  
QY 1031 AAGCAACGCGGCCCGCATGGACAGCAGTGGGCGCAAGAGGAGGCGCGACAGCCCTT 972  
Db 543 SerProGlySerProGlyProAspGly-----LysThrGlyProPro----- 556  
QY 971 CTGCTGCGCTCGGTGGCGCCAGCG-----CTGCTCTCTCAGCCACCA 930  
Db 557 -----GlyProAlaGlyGlnAspGlyArgProGlyProProGlyProPro 571  
QY 929 -----GCAGTGTGCTG 918  
Db 572 GlyAlaAArgGlyGlnAlaGlyValMetGlyPheProGlyProLysGlyAlaAla----- 589  
QY 917 CTACGAGGTGAGGAAGATGAGGTGAGCAGCGCCAAAGAGGCACT-----CCTCTGGG 864  
Db 590 -----GlyGluProGlyLysAlaGlyGluAArgGlyValProGlyProProGly 605  
QY 863 TGCCCAAGGT-- 852  
Db 606 AlaValGlyProAlaGlyLysAspGlyGluAlaGlyAlaGlnGlyProProGlyProAla 625  
QY 851 GGGCAGGCGCACTGGTGTCCCATCTCAATGGCAGCGCAGGA----- 813  
Db 626 GlyProAlaGlyGluAArgGlyGluGlnGlyProAlaGlySerProGlyPheGlnGlyLeu 645  
QY 812 ---GGTAGCCAGCAGCCCGCCCAAGACTGATCATGAAGCATAGACAGAGTAGCCCTGCG 756  
Db 646 ProGlyProAlaGlyProProGlyGluAlaGlyLysProGlyGluGlnGlyValProGly 665  
QY 755 GACAGTGGT---CCGGTCCC-----GGAAGAGGTCCAGAGCAGCGGCTCCAGTGGAG 705  
Db 666 AspLeuGlyAlaProGlyProSerGlyAlaAArgGlyGluArg---GlyPheProGlyGlu 684  
QY 704 TGAAGCACCTGGCCACAGAAAGTCCAGCAGCCCGCCAGGAGTGCAGTGGCA--- 648  
Db 685 ArgGlyValGlnGlyProProGlyProAlaGlyProArgGlyAlaAsnGlyAlaProGly 704  
QY 647 -----GCTCCAGGCGCTGGGATCCGGCAGCAGCAGCC 615  
Db 705 AsnAspGlyAlaLysGlyAspAlaGlyAlaProGlyAla-----ProGlySerGlnGly 722  
QY 614 CTGCTAGCAGCGCGGCTTGGATGAGAAAGCTCAGCAGGATGCCCAAGGACAGTG 555  
Db 723 AlaProGlyLeuGlnGlyMetProGlyGluAArgGlyAlaAlaGlyLeuProGly----- 740  
QY 554 CCCAGATGAAGGCGCGCGGCCCATAGGTCACCGCCAGCGGCTGCTGAGCCCTA 495  
Db 741 ProLysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly-----SerPro 757  
QY 494 GGAGCG---GGACACAGA-----CCAGGCGCCAGCAGCTGGAC 462  
Db 758 GlyLysAspGlyValAArgGlyLeuThrGlyProLysGlyProGlyProAla----- 775  
QY 461 CAATGCCCA----- 453  
Db 776 GlyAlaProGlyAspLysGlyGluSerGlyProSerGlyProAlaGlyProThrGlyAla 795  
QY 452 -----GCACCATGGTCATGAATCTCTCTACCCCACTTCCAGCAGCAGCGGCA 399  
Db 796 ArgGlyAlaProGlyAspArgGlyGluProGlyProProGlyProAlaGlyPheAlaGly 815  
QY 398 CATAGTGTAGCTGCGGCCAAACACACTCCAGCG-----CAAAGTTAGCA 351  
Db 816 -----ProGlyAlaAspGlyGlnProGlyAlaLys 826



```

QY 1343 CTGCTCAGTGGCGAATCGCTGCACACGCGGTCCATGACAGAG-----AGAAGACCA 1290
Db      |||      |||      |||      |||      |||      |||      |||      |||
454 -----Val-GlyValGlnGlyProProGlyProAlaGlyGluGlyLysArgGly 470
QY 1289 GGGAGATGGCGCACTCCAGGACAGCC-----CCAGGCTGCCATCCGAAGCC-----1242
Db      |||      |||      |||      |||      |||      |||      |||      |||
471 AlaArgGlyGluProGlyProThrGlyLeuProGlyProGlyProGlyGluArgGlyGlyPro 490
QY 1241 -----CTTCATCATAGTCTCCGGGCTCGGTGGCGGCTCAGCTC 1200
Db      |||      |||      |||      |||      |||      |||      |||      |||
491 GlySerArgGlyPheProGlyAlaAspGlyValAlaGlyProLysGlyProAlaGlyGlu 510
QY 1199 TGGGCAAGCGCTGTACAGCCCTCGCCACCAATCCGTGTAAACAGCGGTGAAGGTCA 1140
Db      |||      |||      |||      |||      |||      |||      |||      |||
511 ArgGlySerProGlyProAlaGlyProLysGlySerPro-----523
QY 1139 TGAGTCCCATCCAGCTGCACAGCTCAGCCACGAGCGCGGAGGTGGGGGCATGC 1080
Db      |||      |||      |||      |||      |||      |||      |||      |||
523 -----523
QY 1079 GGCAGCACAGCTGGTGCACCGCGGGAAGCAGGCGCCCGCAGGT-----TCCGGA 1032
Db      |||      |||      |||      |||      |||      |||      |||      |||
524 ---GlyGluAlaGlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGly 542
QY 1031 AAGCCAAAGCGCGCGCATGCACAGCAGTGGGGCGACAGAGGGGGCGGACAGCCCTT 972
Db      |||      |||      |||      |||      |||      |||      |||      |||
543 SerProGlySerProGlyProAspGly-----LysThrGlyProPro-----556
QY 971 CTGCTGGCTCGGTGGGGCGCGCG-----CTGCTCTCCAGCCACCA 930
Db      |||      |||      |||      |||      |||      |||      |||      |||
557 -----GlyProAlaGlyGlnAspGlyArgProGlyProProGlyProPro 571
QY 929 -----GCAGTGTGGCTG 918
Db      |||      |||      |||      |||      |||      |||      |||      |||
572 GlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyProLysGlyAlaAla-----589
QY 917 CTACGCAGGTGAGGAAGATGAGGTGAGCAGGCGCAAGAGGACACT-----CCTCCTGG 864
Db      |||      |||      |||      |||      |||      |||      |||      |||
590 -----GlyGluProGlyLysAlaGlyGluArgGlyValProGlyProProGly 605
QY 863 TGCCCAAGT-----AGG 852
Db      |||      |||      |||      |||      |||      |||      |||      |||
606 AlaValGlyProAlaGlyLysAspGlyGluAlaGlyAlaGlnGlyProProGlyProAla 625
QY 851 GGGCCAGGCACTGTGTCTCCAGTCAATGGCAGGAGCA-----813
Db      |||      |||      |||      |||      |||      |||      |||      |||
626 GlyProAlaGlyGluArgGlyGlnGlyProAlaGlySerProGlyPheGlnGlyLeu 645
QY 812 ---GGTAGCCAGGCGAGCGCCCAAGCACTGATCATCAAGGCATAGACAGTAGCCTGGC 756
Db      |||      |||      |||      |||      |||      |||      |||      |||
646 ProGlyProAlaGlyProProGlyGluAlaGlyLysProGlyGluGlnGlyValProGly 665
QY 755 GACAGTGGT---CCGGGTCCC-----GGAAGAGTGCAGAGGAGCGCTCCAGTGGAG 705
Db      |||      |||      |||      |||      |||      |||      |||      |||
666 AspLeuGlyAlaProGlyProSerGlyAlaArgGlyGluArg---GlyPheProGlyGlu 684
QY 704 TGAAGCACACCTGGCCACAGAGTCCAGCGCCCGCCAGCCAGGATGAGCAGTGCCTCA---648
Db      |||      |||      |||      |||      |||      |||      |||      |||
685 ArgGlyValGlnGlyProProGlyProAlaGlyProArgGlyAlaAsnGlyAlaProGly 704
QY 647 -----GCTCCAGGCGCTGGATCCGGGCGACAGCAGCGCC 615
Db      |||      |||      |||      |||      |||      |||      |||      |||
705 AsnAspGlyAlaLysGlyAspAlaGlyAlaProGlyAla-----ProGlySerGlnGly 722
QY 614 CTGCTAGCCAGCGCGCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGACAGTG 555
Db      |||      |||      |||      |||      |||      |||      |||      |||
723 AlaProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuProGly-----740
QY 554 CCCAGATGAAGCGCGCGCGGCGCATAGCTCCAGCGCAGTGTGCTACTGCTCAGCCTA 495
Db      |||      |||      |||      |||      |||      |||      |||      |||
741 ProlysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly-----SerPro 757

```

```

QY 494 GGAGCG-----GGACACAGA-----CCAGGCCCGAGCACTGGAC 462
Db      |||      |||      |||      |||      |||      |||      |||      |||
758 GlyLysAspGlyValArgGlyLeuThrGlyProIleGlyProGlyProAla-----775
QY 461 CAATGCCCA-----453
Db      |||      |||      |||      |||      |||      |||      |||      |||
776 GlyAlaProGlyAspLysGlyGluSerGlyProSerGlyProAlaGlyProThrGlyAla 795
QY 452 -----GCACCATGTCATGAACCTCTCTCTACCCCGCACTTCCAGCAGCAGCGCGCA 399
Db      |||      |||      |||      |||      |||      |||      |||      |||
796 ArgGlyAlaProGlyAspArgGlyGluProGlyProGlyProAlaGlyPheAlaGly 815
QY 398 CATAGTGCATGCTCGCGCCCAACACACACCTCCAGCG-----CAAAGTTAGCA 351
Db      |||      |||      |||      |||      |||      |||      |||      |||
816 -----ProGlyAlaAspGlyGlnProGlyAlaLys 826
QY 350 GGTTGACCAGCAGCTGGCTTTCGGTTCGCCGACGAGCGGCTCACCCAGCAGCTCT 291
Db      |||      |||      |||      |||      |||      |||      |||      |||
827 GlyGluProGlyAspAlaGlyAlaLysGlyAspAla-----GlyProGlyProAla 844
QY 290 GGACCATAGTGGCGCA-----GGCGGTAGGCTCAGGGGCGCTTCAGGCAC 243
Db      |||      |||      |||      |||      |||      |||      |||      |||
845 GlyPro-AlaGlyProGlyProIleGlyAsnValGlyAlaProGlyAlaLysGlyAl 864
QY 242 TCCAGACTGCTGCTCGCTCTGCTCTGCT-----CCAGAGCTGCGGCTCTCTCCTCTG 189
Db      |||      |||      |||      |||      |||      |||      |||      |||
864 a-----ArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAl 879
QY 188 CTGCGCGCACTGCTAGGAATCAGCAGCGCGGCGCTTCTGCCAGCGCTTGTGCGCG 129
Db      |||      |||      |||      |||      |||      |||      |||      |||
879 aAlaGlyArgValGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProG 899
QY 128 TCCAGCTTCTCAGCCCATGCTCAACACACTGCTGTGTGG-----GCACCTCAGTGGG 75
Db      |||      |||      |||      |||      |||      |||      |||      |||
899 yProAlaGlyLysGlu-----GlyGlyLysGlyProArg-GlyG 912
QY 74 ACAGTCTCATCACTCAGATCCTGCGCGCA-----C 45
Db      |||      |||      |||      |||      |||      |||      |||      |||
912 LuThrGlyProAlaGlyArgProGlyGluValGlyProGlyProGlyProGlyProAlaG 932
QY 44 GCGCGCGGCTGCTACCCGCGAGCC-22
Db      |||      |||      |||      |||      |||      |||      |||      |||
932 LyGluLysGlySerProGlyAla 939

```

RESULT 41  
 US-10-096-986-103  
 : Sequence 103, Application US/10096986  
 : Publication No. US20030083464A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Ferrari, Franco A.  
 : Richardson, Charles  
 : Chambers, James  
 : Causey, Stuart  
 : Pollock, Thomas J.  
 : Cappelletto, Joseph  
 : Crissman, John W.  
 : TITLE OF INVENTION: NO. US20030083464A1 Peptides Comprising Repetitive  
 : Units of Amino Acids and DNA Sequences Encoding the Same  
 : NUMBER OF SEQUENCES: 117  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP  
 : STREET: Four Embarcadero Center, Suite 3400  
 : CITY: San Francisco  
 : STATE: California  
 : COUNTRY: US  
 : ZIP: 94111  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/10/096,986



FILING DATE: 12-Mar-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/444,791  
 FILING DATE: 22-Mar-2002  
 APPLICATION NUMBER: US/08/482,085  
 FILING DATE: 07-Jun-1995  
 APPLICATION NUMBER: US/08/175,155  
 FILING DATE: 29-Dec-1993  
 APPLICATION NUMBER: US/08/053,049  
 FILING DATE: 22-Apr-1993  
 APPLICATION NUMBER: US/07/114,618  
 FILING DATE: 29-Oct-1987  
 APPLICATION NUMBER: US/06/927,258  
 FILING DATE: 04-Nov-1986  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Trecartin, Richard F.  
 REGISTRATION NUMBER: 31,801  
 REFERENCE/DOCKET NUMBER: A-55186-11/RFT/BTC  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-781-1989  
 TELEFAX: 415-398-3249  
 INFORMATION FOR SEQ ID NO: 103:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 837 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
 US-10-096-986-103

Alignment Scores:  
 Pred. No.: 1,83e-10 Length: 837  
 Score: 323.00 Matches: 280  
 Percent Similarity: 30.52% Conservative: 45  
 Best Local Similarity: 26.29% Mismatches: 354  
 Query Match: 5.03% Indels: 386  
 DB: 9 Gaps: 59

US-09-759-143-110 (1-3410) x US-10-096-986-103 (1-837)

84 GGTGCCCCACAGCAGAGTGTG-----AGCATGGCTGAGAGTGGACCGGACCA 137  
 37 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 56  
 138 AAGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGCAGTGGCGCAGCAAGAGGA 197  
 57 GlnGly-----LeuProGly----- 61  
 198 GAGGCGGACGCTTCTGGAGCAGACCGGAGCAGAGTCTTGGAGTGGCTGAACGGGCC 257  
 62 -----SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 78  
 258 CTTGAGCCCTACCGCTGGCCCATATGTT-----CCAGAGCTGTGGTGA 305  
 79 GlyAlaProGlyThrProGlyPro-GlnGlyLeuProGlySerProGlyAlaProGlyTh 98  
 306 GCGC---CCTGCTGGCAGCAGCAGCAGCAGTCTTGTGCTGAACCTGCTAACCTTTG 362  
 98 rProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly----- 115  
 363 GCTGAGGCTGTGTTGGCGGAGGAGCATCACCTATGTGCGGCTCTGCTGGTGAAGT-- 420  
 116 -ProGlnGlyLeuProGly-----SerProGlyAlaProGlyThrPr 129  
 421 -GGGGTAGAGAGAGTTCATGACCATGGTGGGCAATGGTCCAGCTGTGGGCT-- 477  
 129 oGlyProGlnGlyLeuProGlySerProGlyAlaPro-----GlyThrProGlyPro 147  
 478 -GCTCTGTGCTCCGCTCCTAGGCTCAGCCAGTGGCGTGGAGCCTATGGCGGCC 536  
 147 nGlyLeuProGlySerPro----- 153

537 GCCGGCCCTCATCTGGCAGCTGTCTCTGGCATCTCTCTGAGCTCTTTCTCATCCAA 596  
 154 -----GlyAlaProGlyThrPro-----GlyProGly 162  
 597 GGGCGCGTGGCTAGCAGGGTGTCT-----GTGCCCCGATCCACAGGCC---CCTGGAGC 647  
 162 nGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySe 182  
 648 TGGCACTGCTCATCTGGCGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 707  
 182 rProGlyAlaProGlyThrProGlyProGlnGlyLeu-----ProGlySerProGlyAl 200  
 708 CACTGGAGGCGCTGCTCTCTGACCTCTTCCGGGA---CCCGGA---CCACTGTCCAGG 761  
 200 aProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 220  
 762 CCTACTCTCTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 821  
 220 y-----ProGlnGlyLeuProGlySerProGlyAla--- 230  
 822 CCATTGACTGGGACACAGTGGCGCTTGGCGCC-----CTACCTGGGACACCCAGGA--- 870  
 231 -----ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPr 246  
 871 -----GGAGTGGCTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911  
 246 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPr 266  
 912 GCGTAGCAGCAGCTGCTGGTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 971  
 266 oGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 286  
 972 AAGGCTGTGCGGCCCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031  
 286 ySerProGlyAlaPro-----GlyThrProGlyProGlnGlyLe 299  
 1032 TCCG---GAACCTGGGCGCTCTGCTTCCCGGCTGACAGCTGCTGCTGCTGCTGCTGCT 1088  
 299 uProGlySerProGlyAlaProGlyThrPro-----GlyProGlnGlyLeuProGlySer 318  
 1089 GCACCTGCTGCGCGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148  
 318 o---GlyAlaProGlyThrProGly----- 325  
 1149 CGCTGTTTACAGCATTCGTTGGCGAGGCGGTGTACAGGCGGTGCGGAGCTGAGC 1208  
 326 -----ProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro----- 339  
 1209 CGGCGACCGGCGGCGGAGACATATGATGAGGCGGTTCGATGGGCGGCTGGGCT-- 1266  
 340 -----GlyProGlnGlyLeu-----ProGlySerProGlyAlaPr 351  
 1267 ---GTTCTCTGAGTGGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322  
 351 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly-- 370  
 1323 AGCATTGCGGACATGAGCAGTCTATTTGGCAGTGTGGCAGTGTGGCAGTGTGGCAGT 1382  
 371 -----ProGlnGlyLeuProGlySerProGlyAla---Pro---GlyThrPro 384  
 1383 GTGCCACATGCTGCTCC-----ACAGTGTGGCGTGGTGGAG 1421  
 385 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro-----Gln 402  
 1422 CTTGAGCGGCTCAGCGGTTACCTTCTCAGCCCTGAGATCCTGCTGCTGCTGCTGCTGCT 1481  
 403 GlyLeuProGlySerProGlyAlaProGly-ThrProGlyPro-----GlnGly 418  
 1482 CCTCCTCTACACCGGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1535  
 418 yLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuPro-GlySerP 438







```
Db 1 ValAlaGluGluAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSer 20
QY 992 TTGTGCGCCCACTGCTGTCCATGCGGGCCCGCTTGGCTTTCCGGAACCTGGGGCCCTG 1051
Db 21 LeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeu 40
QY 1052 CTTCCCGGGTGCACCACTGCTGTCCGCGATGCCCGCACCTGGCGCGCTC 1105
Db 41 LeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgLeu 58

RESULT 44
US-09-895-793-547
; Sequence 547, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 547
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-547
```

```
Alignment Scores:
Pred. No.: 8.01e-10 Length: 58
Score: 312.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.86% Indels: 0
DB: 9 Gaps: 0

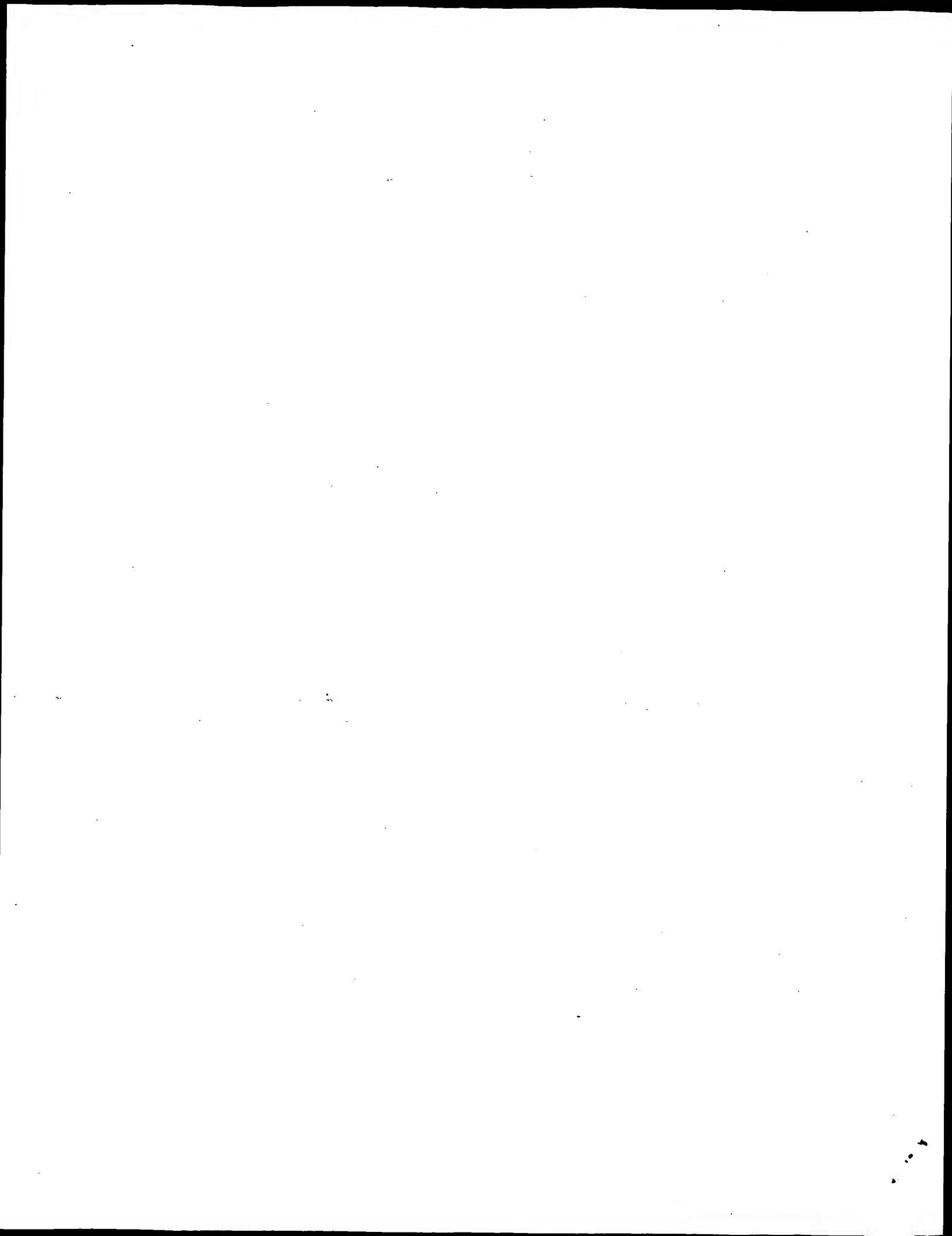
US-09-759-143-110 (1-3410) x US-09-895-793-547 (1-58)
QY 992 GTGGCTGAGGAGCAGCGCTGGGCCCCACCGAGCCAGAGGCGTGTGGCCCCCTCC 991
Db 1 ValAlaGluGluAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSer 20
QY 992 TTGTGCGCCCACTGCTGTCCATGCGGGCCCGCTTGGCTTTCCGGAACCTGGGGCCCTG 1051
Db 21 LeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeu 40
QY 1052 CTTCCCGGGTGCACCACTGCTGTCCGCGATGCCCGCACCTGGCGCGCTC 1105
Db 41 LeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgLeu 58
```

```
RESULT 45
US-09-895-814-547
; Sequence 547, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 547
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-547

Alignment Scores:
Pred. No.: 8.01e-10 Length: 58
Score: 312.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.86% Indels: 0
DB: 9 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-895-814-547 (1-58)
QY 992 GTGGCTGAGGAGCAGCGCTGGGCCCCACCGAGCCAGAGGCGTGTGGCCCCCTCC 991
Db 1 ValAlaGluGluAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSer 20
QY 992 TTGTGCGCCCACTGCTGTCCATGCGGGCCCGCTTGGCTTTCCGGAACCTGGGGCCCTG 1051
Db 21 LeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeu 40
QY 1052 CTTCCCGGGTGCACCACTGCTGTCCGCGATGCCCGCACCTGGCGCGCTC 1105
Db 41 LeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgLeu 58
```

Search completed: June 9, 2003, 22:57:11  
Job time : 249 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 18:34:24 ; Search time 442 Seconds  
(without alignments)  
10797.336 Million cell updates/sec

Title: US-09-759-143-110  
Perfect score: 3410  
Sequence: 1 ggggaacgctgcacgcgc.....aaaaataaaaaaaaaa 3410

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext. 1.0

Searched: 870385 seqs, 699768693 residues  
Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PTCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PTCT\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3409.6	100.0	3410	9	US-09-232-880-110
2	3409.6	100.0	3410	9	US-10-012-896-110
3	3409.6	100.0	3410	9	US-09-895-793-110
4	3409.6	100.0	3410	9	US-09-895-814-110
5	3409.6	100.0	3410	9	US-10-010-940-110
6	3409.6	100.0	3410	10	US-09-745-288-100
7	3409.6	100.0	3410	10	US-09-759-143-110
8	3409.6	100.0	3410	10	US-09-780-669-110
9	3409.6	100.0	3410	10	US-09-030-606-110
10	3409.6	100.0	3410	10	US-09-822-827-110
11	3409.6	100.0	3410	10	US-09-115-453-110
12	3292.4	96.6	3320	10	US-09-838-785-1
13	2585.4	75.8	4034	9	US-10-012-896-704
14	2585.4	75.8	4034	9	US-09-895-793-704
15	2585.4	75.8	4034	9	US-09-895-814-704
16	2585.4	75.8	4034	10	US-09-759-143-704
17	2585.4	75.8	4034	10	US-09-780-669-704
18	2585.4	75.8	4034	10	US-09-822-827-704
19	2196.4	64.4	2904	9	US-10-012-896-703

20	2196.4	64.4	2904	9	US-09-895-793-703	Sequence 703, App
21	2196.4	64.4	2904	9	US-09-895-814-703	Sequence 703, App
22	2196.4	64.4	2904	10	US-09-759-143-703	Sequence 703, App
23	2196.4	64.4	2904	10	US-09-780-669-703	Sequence 703, App
24	2196.4	64.4	2904	10	US-09-822-827-703	Sequence 702, App
25	2142.8	62.8	4894	9	US-10-012-896-702	Sequence 702, App
26	2142.8	62.8	4894	9	US-09-895-793-702	Sequence 702, App
27	2142.8	62.8	4894	9	US-09-895-814-702	Sequence 702, App
28	2142.8	62.8	4894	10	US-09-759-143-702	Sequence 702, App
29	2142.8	62.8	4894	10	US-09-780-669-702	Sequence 702, App
30	2142.8	62.8	4894	10	US-09-822-827-702	Sequence 705, App
31	1815.8	53.2	6976	9	US-10-012-896-705	Sequence 705, App
32	1815.8	53.2	6976	9	US-09-895-793-705	Sequence 705, App
33	1815.8	53.2	6976	9	US-09-895-814-705	Sequence 705, App
34	1815.8	53.2	6976	10	US-09-759-143-705	Sequence 705, App
35	1815.8	53.2	6976	10	US-09-780-669-705	Sequence 705, App
36	1815.8	53.2	6976	10	US-09-822-827-705	Sequence 851, App
37	794.6	23.3	1203	9	US-10-012-896-851	Sequence 851, App
38	794.6	23.3	1203	9	US-09-895-793-851	Sequence 851, App
39	794.6	23.3	1203	9	US-09-895-814-851	Sequence 851, App
40	794.6	23.3	1203	10	US-09-759-143-851	Sequence 851, App
41	794.6	23.3	1203	10	US-09-780-669-851	Sequence 851, App
42	794.6	23.3	1203	10	US-09-822-827-851	Sequence 10, Appl
43	673.4	19.7	789	9	US-09-232-880-10	Sequence 10, Appl
44	673.4	19.7	789	9	US-10-012-896-10	Sequence 10, Appl
45	673.4	19.7	789	9	US-09-895-793-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-232-880-110  
; Sequence 110, Application US/09232880  
; Publication No. US20020182596A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer Lynn  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF  
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.428C6  
; CURRENT APPLICATION NUMBER: US/09/232,880  
; CURRENT FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 338  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 110  
; LENGTH: 3410  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-232-880-110

Query Match	100.0%	Score	3409.6	DB	9	Length	3410
Best Local Similarity	100.0%	Pred. No.	0				
Matches	3410	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	GGGAACACGCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCTCGGCCAGGATCTGA	60				
DB	1	GGGAACACGCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCTCGGCCAGGATCTGA	60				
QY	61	GTGATGAGACGTGTCTCCCACTGAGTGGCCCAACAGCAGGCTGTGAGCATGGCTGAG	120				
DB	61	GTGATGAGACGTGTCTCCCACTGAGTGGCCCAACAGCAGGCTGTGAGCATGGCTGAG	120				
QY	121	AAGCTGGACCGCCACCAAGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGGAGTT	180				
DB	121	AAGCTGGACCGCCACCAAGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGGAGTT	180				
QY	181	GGCGGACGACGAGGAGGAGCGCGAGCTTCTTGAGCAGCGCGGAGCAAGCAGTTCTG	240				
DB	181	GGCGGACGACGAGGAGGAGCGCGAGCTTCTTGAGCAGCGCGGAGCAAGCAGTTCTG	240				
QY	241	GAGTGCCTGAACGGGCCCCCTGAGCCCTACCCGCTGGCCCACTATGTCAGGAGCTGTG	300				



Db 241 GAGTGCCTGAACGGCCCTGAGCCCTACCGCCTGGCCACATATGTCACAGGCTGTG 300  
QY 301 GGTGAGCGCGCTGTGGGACACCGAAGCCAGCTCTTGGTGAACCTGTACCTTT 360  
Db 301 GGTGAGCGCGCTGTGGGACACCGAAGCCAGCTCTTGGTGAACCTGTACCTTT 360  
QY 361 TGGCCTGGAGGTGTGTTTGGCGGAGGATCACCTATGTGCGCGCTGTGCTGGAAGT 420  
Db 361 TGGCCTGGAGGTGTGTTTGGCGGAGGATCACCTATGTGCGCGCTGTGCTGGAAGT 420  
QY 421 GGGGGTAGAGGAGAGTTTATGACCATGTGCTGGGCAATGGTCCAGTGTGGCCTGT 480  
Db 421 GGGGGTAGAGGAGAGTTTATGACCATGTGCTGGGCAATGGTCCAGTGTGGCCTGT 480  
QY 481 CTGTGTCCCGCTCTAGGCTCAGCCAGTACCACTGGCGTGGACGCTATGGCGCGCGG 540  
Db 481 CTGTGTCCCGCTCTAGGCTCAGCCAGTACCACTGGCGTGGACGCTATGGCGCGCGG 540  
QY 541 GCCCTTCATCTGGGCACTGTCTTGGGCAATCTGTGCTGAGCCTCTTCTCATCCCAAGGC 600  
Db 541 GCCCTTCATCTGGGCACTGTCTTGGGCAATCTGTGCTGAGCCTCTTCTCATCCCAAGGC 600  
QY 601 CGGCTGTGCTAGCAGGCTGTGTCGCGGATCCAGGCGCTGGAGCTGGCACTGCTCAT 660  
Db 601 CGGCTGTGCTAGCAGGCTGTGTCGCGGATCCAGGCGCTGGAGCTGGCACTGCTCAT 660  
QY 661 CTTGGGCGTGGGCTGTGCTGAGCTTCTGTGGCCAGTGTGCTTCACTCCACTGGAGCCT 720  
Db 661 CTTGGGCGTGGGCTGTGCTGAGCTTCTGTGGCCAGTGTGCTTCACTCCACTGGAGCCT 720  
QY 721 GCTCTCTGACCTCTCCGGGACCCGACCACTGTGCCAGGCTACTGTCTATGCTTT 780  
Db 721 GCTCTCTGACCTCTCCGGGACCCGACCACTGTGCCAGGCTACTGTCTATGCTTT 780  
QY 781 CATGATCAGTCTTGGGGCTGTGCTGGCTACCTCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 CATGATCAGTCTTGGGGCTGTGCTGGCTACCTCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 TGCCCTGGCCCTACTTGGGACCCAGGAGAGTGCCTTGTGGCTGCTGCTGCTGCTCAT 900  
Db 841 TGCCCTGGCCCTACTTGGGACCCAGGAGAGTGCCTTGTGGCTGCTGCTGCTGCTCAT 900  
QY 901 CTTCTCTACCTGCTAGCAGCCACACTGTGCTGTGCTGAGGAGGAGGCTGCTTGGCTGCT 960  
Db 901 CTTCTCTACCTGCTAGCAGCCACACTGTGCTGTGCTGAGGAGGAGGCTGCTTGGCTGCT 960  
QY 961 CGAGCAGCAGAGGCTGTGCGGCCCTCTTGTGCGCCCACTGCTGCTGCTGCTGCTGCT 1020  
Db 961 CGAGCAGCAGAGGCTGTGCGGCCCTCTTGTGCGCCCACTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 CGGCTTGGCTTTCGGAACTGGCGCCCTGCTTCCCGGCTGCAACAGCTGTGCTGCGG 1080  
Db 1021 CGGCTTGGCTTTCGGAACTGGCGCCCTGCTTCCCGGCTGCAACAGCTGTGCTGCGG 1080  
QY 1081 CATGCCCGCACCTCGCGGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 1081 CATGCCCGCACCTCGCGGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
QY 1141 GACCTTACGCTGTTTTACAGGATTTGCTGGGCGAGGGGCTGTACAGGCGCTGCCAG 1200  
Db 1141 GACCTTACGCTGTTTTACAGGATTTGCTGGGCGAGGGGCTGTACAGGCGCTGCCAG 1200  
QY 1201 AGCTGAGCGGCGACCGAGGCGCGGACACTATGATGAAGCGCTTCGGATGGGAGCCT 1260  
Db 1201 AGCTGAGCGGCGACCGAGGCGCGGAGACTATGATGAAGCGCTTCGGATGGGAGCCT 1260  
QY 1261 GGGGCTGTTCCTGAGTGCCTCATCTCCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
Db 1261 GGGGCTGTTCCTGAGTGCCTCATCTCCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 CGAGGATTCGCACTCAGCAGTCTATTTGGCAGTGTGCGAGCTTTCCTGTGCTGCTGCT 1380  
Db 1321 CGAGGATTCGCACTCAGCAGTCTATTTGGCAGTGTGCGAGCTTTCCTGTGCTGCTGCT

Db 1321 GCAGCAATTCGCACTCAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCTGTGCTGCTGCT 1380  
QY 1381 CGGTGCCACATGCTTGTCCACAGTGTGGCCGTGTGACAGCTTTCAGCGCGCTCACCGG 1440  
Db 1381 CGGTGCCACATGCTTGTCCACAGTGTGGCCGTGTGACAGCTTTCAGCGCGCTCACCGG 1440  
QY 1441 GTTCACTTCTCAGCGCTGACATGCTGCCCTTACACACTGGCCTCCCTTACACCGGGA 1500  
Db 1441 GTTCACTTCTCAGCGCTGACATGCTGCCCTTACACACTGGCCTCCCTTACACCGGGA 1500  
QY 1501 GAAGCAGTGTTCCTGCCCAATACCGAGGGGACACTGGAGTGTACAGTGTAGGACAG 1560  
Db 1501 GAAGCAGTGTTCCTGCCCAATACCGAGGGGACACTGGAGTGTACAGTGTAGGACAG 1560  
QY 1561 CTTGATGACAGCTTCTTCCGAGGCGCTTAAGCCTGGAGCTCCCTTCCCTTAATGACAGCT 1620  
Db 1561 CTTGATGACAGCTTCTTCCGAGGCGCTTAAGCCTGGAGCTCCCTTCCCTTAATGACAGCT 1620  
QY 1621 GGGTCTGGAGCAGTGGCCTGCTCCACCTTCCACCGCGCTCTGCGGGGCTCTGCGCTG 1680  
Db 1621 GGGTCTGGAGCAGTGGCCTGCTCCACCTTCCACCGCGCTCTGCGGGGCTCTGCGCTG 1680  
QY 1681 TGATCTCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
Db 1681 TGATCTCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
QY 1741 GGGCATCTCGCTGGACCTCGCCATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1800  
Db 1741 GGGCATCTCGCTGGACCTCGCCATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1800  
QY 1801 ATCCCTCTTTATGGCTCCATTTGCCAGCTCAGCCAGTCTGCTGCTGCTGCTGCTGCTGCT 1860  
Db 1801 ATCCCTCTTTATGGCTCCATTTGCCAGCTCAGCCAGTCTGCTGCTGCTGCTGCTGCTGCT 1860  
QY 1861 TGCCGCGGCGCTGGTCTGGTGGCCTTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
Db 1861 TGCCGCGGCGCTGGTCTGGTGGCCTTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
QY 1921 CGACTTGGCCAAATACTCAGCTAGAAACTTCCAGCACATTTGGGGTGGAGGCGCTGCT 1980  
Db 1921 CGACTTGGCCAAATACTCAGCTAGAAACTTCCAGCACATTTGGGGTGGAGGCGCTGCT 1980  
QY 1981 CACTTGGTTCAGCTTCCCGCTTCTGTTAGCCCATGGGCTGCGCGGCTGCGCGGCTGCT 2040  
Db 1981 CACTTGGTTCAGCTTCCCGCTTCTGTTAGCCCATGGGCTGCGCGGCTGCGCGGCTGCT 2040  
QY 2041 TTCTGTTGCTGCCAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
Db 2041 TTCTGTTGCTGCCAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
QY 2101 GCTGACAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
Db 2101 GCTGACAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
QY 2161 ACTGAGGCGCTTCCAGGGGTTTACGTGAGCTTATACAGGAGGCGCAAGAGGCTGCT 2220  
Db 2161 ACTGAGGCGCTTCCAGGGGTTTACGTGAGCTTATACAGGAGGCGCAAGAGGCTGCT 2220  
QY 2221 ATGCACTGGAATCGGGGCTGCTGAGTGGTAACTTCCAGGCTCAGGCTTAAACAGTAGC 2280  
Db 2221 ATGCACTGGAATCGGGGCTGCTGAGTGGTAACTTCCAGGCTCAGGCTTAAACAGTAGC 2280  
QY 2281 CTCTAGTTGAGACACACTTAGAAGGGTTTTTGGAGCTGAATAAAGCTCAGTACCTG 2340  
Db 2281 CTCTAGTTGAGACACACTTAGAAGGGTTTTTGGAGCTGAATAAAGCTCAGTACCTG 2340  
QY 2341 GTTTCCTCATCTTAAGCGCTTAACTGAGCTTCTGTTTAACTGAGCTTCTGATGGGAG 2400  
Db 2341 GTTTCCTCATCTTAAGCGCTTAACTGAGCTTCTGTTTAACTGAGCTTCTGATGGGAG 2400  
QY 2401 TTTCTAGGATGAACACTCTCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460  
Db 2401 TTTCTAGGATGAACACTCTCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460



Db 481 CTGTGTCCTCCCTCCTAGGCTCAGCCAGTACCACTGGCGTGGAGCGTATGGCGCGCGCG 540  
Qy 541 GCCCTTCATCTGGGCACTGTCTTGGGCATCTCTGCTGAGCCTCTTCTCATCCCAAGGC 600  
Db 541 GCCCTTCATCTGGGCACTGTCTTGGGCATCTCTGCTGAGCCTCTTCTCATCCCAAGGC 600  
Qy 601 CGGCTGGCTAGCAGGCTGTCTGTGCCCGATCCAGGCGCTGGAGCTGGCACTGCTCAT 660  
Db 601 CGGCTGGCTAGCAGGCTGTCTGTGCCCGATCCAGGCGCTGGAGCTGGCACTGCTCAT 660  
Qy 661 CTTGGGCGTGGGCTGTCTGGACTTCTGTGCCAGGTGTCTTCACTCCACTGGAGCCCT 720  
Db 661 CTTGGGCGTGGGCTGTCTGGACTTCTGTGCCAGGTGTCTTCACTCCACTGGAGCCCT 720  
Qy 721 GCTCTGTAGCTCTTCCGGACCCGACCACTGTGCCAGGCTACTCTGTATGCCCTT 780  
Db 721 GCTCTGTAGCTCTTCCGGACCCGACCACTGTGCCAGGCTACTCTGTATGCCCTT 780  
Qy 781 CATGATCAGTCTTGGGGCTGGCTGGCTACCTTACCTTGCCTGCCATGACTGGGACACAG 840  
Db 781 CATGATCAGTCTTGGGGCTGGCTGGCTACCTTACCTTGCCTGCCATGACTGGGACACAG 840  
Qy 841 TGCCCTGGCCCTTACCTGGGACCCAGGAGTGCCTTCTTGGCTGCTCACCTCAT 900  
Db 841 TGCCCTGGCCCTTACCTGGGACCCAGGAGTGCCTTCTTGGCTGCTCACCTCAT 900  
Qy 901 CTTCTCAGCTGGGTAGCAGCACACTGTGTGGCTGAGAGGAGCGCTGGGCCCCAC 960  
Db 901 CTTCTCAGCTGGGTAGCAGCACACTGTGTGGCTGAGAGGAGCGCTGGGCCCCAC 960  
Qy 961 CGAGCAGCAGAGGCTGTGGGCCCCCTTGTGGCCCCCTGCTGTCCATGCCGGC 1020  
Db 961 CGAGCAGCAGAGGCTGTGGGCCCCCTTGTGGCCCCCTGCTGTCCATGCCGGC 1020  
Qy 1021 CGCCTTGGCTTTCGGAACTGGGCGCTGCTTCCCGGCTGCACCACTGTGTGCGG 1080  
Db 1021 CGCCTTGGCTTTCGGAACTGGGCGCTGCTTCCCGGCTGCACCACTGTGTGCGG 1080  
Qy 1081 CATGCCCGGACCTTCCGCGGCTGTGTGGCTGAGCTGTGCACTGTGCTGCTGCT 1140  
Db 1081 CATGCCCGGACCTTCCGCGGCTGTGTGGCTGAGCTGTGCACTGTGCTGCTGCT 1140  
Qy 1141 GACCTTCAGCTGTCTTACACGATTTGCTGGCGAGGCTGTACAGGCGTGCACG 1200  
Db 1141 GACCTTCAGCTGTCTTACACGATTTGCTGGCGAGGCTGTACAGGCGTGCACG 1200  
Qy 1201 AGCTAGCGCGGACCGAGGCGCGGAGACTATGATGAAGCGCTTCCGATGGGCGCT 1260  
Db 1201 AGCTAGCGCGGACCGAGGCGCGGAGACTATGATGAAGCGCTTCCGATGGGCGCT 1260  
Qy 1261 GGGGCTGTCTGAGTGGCCATCTCCCTGGTCTTCTCTGTGCTATGGACGGCTGGT 1320  
Db 1261 GGGGCTGTCTGAGTGGCCATCTCCCTGGTCTTCTCTGTGCTATGGACGGCTGGT 1320  
Qy 1321 GCAGGATTCGGCACTCGAGCACTTATTTGGCCAGTGTGGCAGCTTTCCTGTGGCTG 1380  
Db 1321 GCAGGATTCGGCACTCGAGCACTTATTTGGCCAGTGTGGCAGCTTTCCTGTGGCTG 1380  
Qy 1381 CGGTGCCACATGCCCTGCCACAGTGTGGCGTGGTGACAGCTTACGCGCCCTCACCGG 1440  
Db 1381 CGGTGCCACATGCCCTGCCACAGTGTGGCGTGGTGACAGCTTACGCGCCCTCACCGG 1440  
Qy 1441 GTTCACTTCTCAGCCCTGCAGATCTCCCTACACACTGGCCTCCCTCTACACCGGGA 1500  
Db 1441 GTTCACTTCTCAGCCCTGCAGATCTCCCTACACACTGGCCTCCCTCTACACCGGGA 1500  
Qy 1501 GAAGCAGTGTCTTCCGCCAAATACCGAGGAGACACTGGAGGTGCTAGCAGTGGAGACAG 1560  
Db 1501 GAAGCAGTGTCTTCCGCCAAATACCGAGGAGACACTGGAGGTGCTAGCAGTGGAGACAG 1560  
Qy 1561 CTTGATGACCACTTCTTCCAGGCGCTTAAAGCTGGAGCTCCCTTCCCTAATGACACAGT 1620  
Db 1561 CTTGATGACCACTTCTTCCAGGCGCTTAAAGCTGGAGCTCCCTTCCCTAATGACACAGT 1620

Qy 1621 GGGTCTTGAGGCACTGGCTGCTCCACCTCCACCGCGCTCTGCGGGCGCTCTGCGTG 1680  
Db 1621 GGGTCTTGAGGCACTGGCTGCTCCACCTCCACCGCGCTCTGCGGGCGCTCTGCGTG 1680  
Qy 1681 TGATGCTCCGCTAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
Db 1681 TGATGCTCCGCTAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
Qy 1741 GGGCATCTGCCCTGGACCTCGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
Db 1741 GGGCATCTGCCCTGGACCTCGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
Qy 1801 ATCCCTGTTTATGGCTCCATTTGCTCAGCTCAGCCAGTCTGCTGCTGCTGCTGCTGCT 1860  
Db 1801 ATCCCTGTTTATGGCTCCATTTGCTCAGCTCAGCCAGTCTGCTGCTGCTGCTGCTGCT 1860  
Qy 1861 TGCCGAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920  
Db 1861 TGCCGAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920  
Qy 1921 CGACTTGGCCAAATACCTAGCTAGAACTTCCAGCACATTTGGGGTGGAGGGCTGCCT 1980  
Db 1921 CGACTTGGCCAAATACCTAGCTAGAACTTCCAGCACATTTGGGGTGGAGGGCTGCCT 1980  
Qy 1981 CACTGGGTCCAGCTCCCGCTGTTAGCCCCATGGGGTGGGGCTGGCGGCGCAGT 2040  
Db 1981 CACTGGGTCCAGCTCCCGCTGTTAGCCCCATGGGGTGGGGCTGGCGGCGCAGT 2040  
Qy 2041 TTCTGTTGCTGCCAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
Db 2041 TTCTGTTGCTGCCAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
Qy 2101 GCTCAGAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
Db 2101 GCTCAGAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
Qy 2161 ACTGAGGCTTCCAGGGGTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220  
Db 2161 ACTGAGGCTTCCAGGGGTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220  
Qy 2221 ATGCACTTGAATCGGGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Db 2221 ATGCACTTGAATCGGGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Qy 2281 CTCTAGTTGAGACACACTAGAGAGGTTTGGAGGCTGAATAAAGCTGCTGCTGCTG 2340  
Db 2281 CTCTAGTTGAGACACACTAGAGAGGTTTGGAGGCTGAATAAAGCTGCTGCTGCTG 2340  
Qy 2341 GTTTCCTCATCTTAAGCCCTTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
Db 2341 GTTTCCTCATCTTAAGCCCTTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
Qy 2401 TTTCTAGGATGAACACTCTCTCCATGGGATTTGAACATATGACTTATTTAGGGGAAGA 2460  
Db 2401 TTTCTAGGATGAACACTCTCTCCATGGGATTTGAACATATGACTTATTTAGGGGAAGA 2460  
Qy 2461 GTCTGAGGGGCAACACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
Db 2461 GTCTGAGGGGCAACACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
Qy 2521 GATCCACCCCTCTTACCTTTATCAGGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580  
Db 2521 GATCCACCCCTCTTACCTTTATCAGGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580  
Qy 2581 CAGAGACAGGCTTAAATATTTAACTTATTTAACTTATTTAACTTATTTAACTTATTT 2640  
Db 2581 CAGAGACAGGCTTAAATATTTAACTTATTTAACTTATTTAACTTATTTAACTTATTT 2640  
Qy 2641 TGCTAGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
Db 2641 TGCTAGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700





```
QY 3001 CCCAACTTTCCCTACCCCAACTTTCCCAACAGCTCCACAAACCTGTTGGAGCTACT 3060
Db 3001 CCCAACTTTCCCTACCCCAACTTTCCCAACAGCTCCACAAACCTGTTGGAGCTACT 3060
QY 3061 GCAGGACCAAGAGCAAAAGTGGGTTTCCCAAGCTTTTGTCCATCTCAGCCCCCAGAGT 3120
Db 3061 GCAGGACCAAGAGCAAAAGTGGGTTTCCCAAGCTTTTGTCCATCTCAGCCCCCAGAGT 3120
QY 3121 ATATCTGTGTGGGGAATCTCACACAGAAATCAGGACACCCCTGCCTGAGCTAAGG 3180
Db 3121 ATATCTGTGTGGGGAATCTCACACAGAAATCAGGACACCCCTGCCTGAGCTAAGG 3180
QY 3181 GAGGTCTTATCTCTCAGGGGGGTTTAAAGTGGCGTTTGAATAATCTGCTTATTATT 3240
Db 3181 GAGGTCTTATCTCTCAGGGGGGTTTAAAGTGGCGTTTGAATAATCTGCTTATTATT 3240
QY 3241 TAGCGGGGTGAATTTTATCTGTAAGTGAGCAATCAGAGTATATGTTTATGTGACA 3300
Db 3241 TAGCGGGGTGAATTTTATCTGTAAGTGAGCAATCAGAGTATATGTTTATGTGACA 3300
QY 3301 AAATTAAGGCTTTCTTATATGTTTAAAAAATAAATAAATAAATAAATAAATAA 3360
Db 3301 AAATTAAGGCTTTCTTATATGTTTAAAAAATAAATAAATAAATAAATAAATAA 3360
QY 3361 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3410
Db 3361 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3410
```

## RESULT 4

```
US-09-895-814-110
; Sequence 110, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillion, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Barrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Repler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 3410
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-895-814-110
```

```
Query Match 100.0%; Score 3409.6; DB 9; Length 3410;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGAACACAGCCTGCACGCGCTGCTCCGGGTGACAGCCGCGCTCGGCAGGATCTGA 60
Db 1 GGAACACAGCCTGCACGCGCTGCTCCGGGTGACAGCCGCGCTCGGCAGGATCTGA 60
QY 61 GTGATGAGACGTGTGCCCACTGAGGTGCCCAACAGCAGAGGTGTTGAGCATGGCTGAG 120
Db 61 GTGATGAGACGTGTGCCCACTGAGGTGCCCAACAGCAGAGGTGTTGAGCATGGCTGAG 120
QY 121 AAGCTGGACCGGACCAAGAGGCTGGCAGAAATGGCGCCTGCTGATTCCTAGCAGATT 180
Db 121 AAGCTGGACCGGACCAAGAGGCTGGCAGAAATGGCGCCTGCTGATTCCTAGCAGATT 180
QY 181 GCGCAGCAAGAGGAGAGGCGCAGCTTCTGGAGCAGAGCCGAGAGCAAGCAAGTCTTG 240
Db 181 GCGCAGCAAGAGGAGAGGCGCAGCTTCTGGAGCAGAGCCGAGAGCAAGCAAGTCTTG 240
QY 241 GAGTGCCTGAACGGCGCCCTGAGCCCTTACCGCCTGGCCCACTATGTTCCAGAGGCTGTG 300
Db 241 GAGTGCCTGAACGGCGCCCTGAGCCCTTACCGCCTGGCCCACTATGTTCCAGAGGCTGTG 300
QY 301 GGTGAGCGCCTGCTGCGGCAACCGGAAAGCCAGCTCTTGTGTGCTCAACCTGCTAACCTT 360
Db 301 GGTGAGCGCCTGCTGCGGCAACCGGAAAGCCAGCTCTTGTGTGCTCAACCTGCTAACCTT 360
QY 361 TGGCCTGAGAGGTGTTTGGCGCAGGCAATCACCTATGTCGCCCTCTGCTGCTGGAAGT 420
Db 361 TGGCCTGAGAGGTGTTTGGCGCAGGCAATCACCTATGTCGCCCTCTGCTGCTGGAAGT 420
QY 421 GGGGTGAGAGGAGAGTTCATGACCATGGTCTGGGCATTTGGTCAGTGTCTGGGCTGGT 480
Db 421 GGGGTGAGAGGAGAGTTCATGACCATGGTCTGGGCATTTGGTCAGTGTCTGGGCTGGT 480
QY 481 CTGTGTCGCGCTCTAGGCTCAGCAGTGACCACTGGCGTGAGCACTATGGCGCGCGCG 540
Db 481 CTGTGTCGCGCTCTAGGCTCAGCAGTGACCACTGGCGTGAGCACTATGGCGCGCGCG 540
QY 541 GGCCTTCATCTGGGCACCTGCTTGGGCATCTGCTGAGCCTCTTCTTCATCCCAAGGCG 600
Db 541 GGCCTTCATCTGGGCACCTGCTTGGGCATCTGCTGAGCCTCTTCTTCATCCCAAGGCG 600
QY 601 CGGCTGGCTAGCAGGCTGCTGTGCCCGGATCCAGGCGCTGAGCTGGGAGCTGCTCAT 660
Db 601 CGGCTGGCTAGCAGGCTGCTGTGCCCGGATCCAGGCGCTGAGCTGGGAGCTGCTCAT 660
QY 661 CTGGGCGTGGGCTGCTGACTTCTGTGGCCAGGTGTGCTTCACTTCCACTGAGGCGCT 720
Db 661 CTGGGCGTGGGCTGCTGACTTCTGTGGCCAGGTGTGCTTCACTTCCACTGAGGCGCT 720
QY 721 GCTCTCTGACCTTCTCCGGGACCGGACCACTGTGTCGCCAGGCTTACTGTCTATGCTT 780
Db 721 GCTCTCTGACCTTCTCCGGGACCGGACCACTGTGTCGCCAGGCTTACTGTCTATGCTT 780
QY 781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 TGCCCTGGGCGCTTCTGCTGGGACCGGAGGAGTGCCTTCTTGGGCTGCTGCTGCTGCT 900
Db 841 TGCCCTGGGCGCTTCTGCTGGGACCGGAGGAGTGCCTTCTTGGGCTGCTGCTGCTGCT 900
QY 901 CTTCTCTACCTGCGTAGCAGCCACACTGTGTGGTGTGAGGAGGAGGAGGAGGAGGAGG 960
Db 901 CTTCTCTACCTGCGTAGCAGCCACACTGTGTGGTGTGAGGAGGAGGAGGAGGAGGAGG 960
QY 961 CGAGCCAGCAGAGGAGGCTGTGGCGCCCTCTGCTGTCGCCCGCTGCTGCTGCTGCTG 1020
Db 961 CGAGCCAGCAGAGGAGGCTGTGGCGCCCTCTGCTGTCGCCCGCTGCTGCTGCTGCTG 1020
QY 1021 CCCTTTGGCTTTCCGGAACCTGGGCGCCCTGCTGCTGTCGCCCGCTGCTGCTGCTGCTG 1080
Db 1021 CCCTTTGGCTTTCCGGAACCTGGGCGCCCTGCTGCTGTCGCCCGCTGCTGCTGCTGCTG 1080
```



QY 1081 CATGCCCGCCACCTCGCGCGGCTCTTCGTGGCTGAGCTGTGACGTGGATGGCACTCAT 1140  
Db CATGCCCGCCACCTCGCGCGGCTCTTCGTGGCTGAGCTGTGACGTGGATGGCACTCAT 1140  
QY 1141 GACCTTCACGCTGTTTACACGATTTTCGTGGCGAGGGCTCTACACAGGGCTGCCAG 1200  
Db GACCTTCACGCTGTTTACACGATTTTCGTGGCGAGGGCTCTACACAGGGCTGCCAG 1200  
QY 1201 AGCTGAGCGGGACCGAGGCGCCGGAGACACTATGATGAAGGGTTCGGATGGCAGCCT 1260  
Db AGCTGAGCGGGACCGAGGCGCCGGAGACACTATGATGAAGGGTTCGGATGGCAGCCT 1260  
QY 1261 GGGCTGCTCCTGCAGTGGCCATCTCCCTGGTCTCTCTGCTGATGACCGGGTGT 1320  
Db GGGCTGCTCCTGCAGTGGCCATCTCCCTGGTCTCTCTGCTGATGACCGGGTGT 1320  
QY 1321 GCAGCGATTGCGGACCTGAGCAGTCTATTGGCCAGTGTGGCAGCTTCCCTGTGCTGC 1380  
Db GCAGCGATTGCGGACCTGAGCAGTCTATTGGCCAGTGTGGCAGCTTCCCTGTGCTGC 1380  
QY 1381 CGGTGCCATGCTGTCCACAGTGTGGCGGTGGTGAAGCTTCAGCGCCCTCACCGG 1440  
Db CGGTGCCATGCTGTCCACAGTGTGGCGGTGGTGAAGCTTCAGCGCCCTCACCGG 1440  
QY 1441 GTTTCACCTTCTAGCCCTCGAGATCTGCTTACACACTGGCTCCCTTACACCGGGA 1500  
Db GTTTCACCTTCTAGCCCTCGAGATCTGCTTACACACTGGCTCCCTTACACCGGGA 1500  
QY 1501 GAAGCAGGTGTTCTGCCCAATACCGAGGGGACACTGGAGGTGTAGCAGTGAAGACAG 1560  
Db GAAGCAGGTGTTCTGCCCAATACCGAGGGGACACTGGAGGTGTAGCAGTGAAGACAG 1560  
QY 1561 CCTGATGACAGCTTCTGCCAGGCGCTAAGCCTGGAGCTCCCTTCCCTAATGAGACGT 1620  
Db CCTGATGACAGCTTCTGCCAGGCGCTAAGCCTGGAGCTCCCTTCCCTAATGAGACGT 1620  
QY 1621 GGTGCTGGAGGAGTGGCTGTCTCCACCTCCACCGGGCTCTGCGGGGCTCTGCGCTG 1680  
Db GGTGCTGGAGGAGTGGCTGTCTCCACCTCCACCGGGCTCTGCGGGGCTCTGCGCTG 1680  
QY 1681 TGATGCTCCGACGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
Db TGATGCTCCGACGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
QY 1741 GGGCATCTGCTGGACCTCGCCATCTGATAGTGCCTTCCGCTGCTCCAGGTGGCC 1800  
Db GGGCATCTGCTGGACCTCGCCATCTGATAGTGCCTTCCGCTGCTCCAGGTGGCC 1800  
QY 1801 ATCCCTGTTTATGGGCTCCATTTGCTCAGCTCAGCCAGTCTGCTATATGCTGTC 1860  
Db ATCCCTGTTTATGGGCTCCATTTGCTCAGCTCAGCCAGTCTGCTATATGCTGTC 1860  
QY 1861 TGCCGAGGCGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920  
Db TGCCGAGGCGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920  
QY 1921 CGACTTGGCCAAATACTCAGCGTAGAAATTTCCAGCACATTTGGGGTGGAGGCGCTGCT 1980  
Db CGACTTGGCCAAATACTCAGCGTAGAAATTTCCAGCACATTTGGGGTGGAGGCGCTGCT 1980  
QY 1981 CACTGGGTCAGCTCCCGCTCTCTGTTAGCCCCATGGGGTGGCGGCTGGCGCCAGT 2040  
Db CACTGGGTCAGCTCCCGCTCTCTGTTAGCCCCATGGGGTGGCGGCTGGCGCCAGT 2040  
QY 2041 TTCTGTTGCTGCAAGTAATGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
Db TTCTGTTGCTGCAAGTAATGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
QY 2101 GCTGCACAGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGG 2160  
Db GCTGCACAGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGG 2160  
QY 2161 ACTGAGGCGCTTCCAGGGGGTTCAGTCTGAGCTTATACAGGGAGGCGCCAGAGGGCTCC 2220

Db 2161 ACTGAGGCGCTTCCAGGGGGTTCAGTCTGAGCTTATACAGGGAGGCGCCAGAGGGCTCC 2220  
QY 2221 ATGCACTGGAATGCGGGGACTCTGAGGTGGATTAACAGGCTCAGGCTTAACAGCTAGC 2280  
Db ATGCACTGGAATGCGGGGACTCTGAGGTGGATTAACAGGCTCAGGCTTAACAGCTAGC 2280  
QY 2281 CTCTCTAGTTGAGACACACTAGAGAGGGTTCGGGAGCTGAATAAATACTCAGTCACTG 2340  
Db CTCTCTAGTTGAGACACACTAGAGAGGGTTCGGGAGCTGAATAAATACTCAGTCACTG 2340  
QY 2341 GTTCCCATCTCTAAGCCCTTAACCTGCACTGCTGTTTAAATGAGTCTTTCATGGGAG 2400  
Db GTTCCCATCTCTAAGCCCTTAACCTGCACTGCTGTTTAAATGAGTCTTTCATGGGAG 2400  
QY 2401 TTTCTAGATGAACACTCTCCATGGATTTGAACATATGACTTATTTGAGGGGAGA 2460  
Db TTTCTAGATGAACACTCTCCATGGATTTGAACATATGACTTATTTGAGGGGAGA 2460  
QY 2461 GTCTGAGGGGCAACACAGAGTCCCTCAGCCACACAGCTGCTGCTTTTGGCT 2520  
Db GTCTGAGGGGCAACACAGAGTCCCTCAGCCACACAGCTGCTGCTTTTGGCT 2520  
QY 2521 GATCACCCCTCTTACCTTTTATCAGGATGTGGCTGTTGGCTCTTCTTCTGCTGCCATCA 2580  
Db GATCACCCCTCTTACCTTTTATCAGGATGTGGCTGTTGGCTCTTCTTCTGCTGCCATCA 2580  
QY 2581 CAGACACAGGCAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATA 2640  
Db CAGACACAGGCAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATA 2640  
QY 2641 TGCTAGCTTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2700  
Db TGCTAGCTTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2700  
QY 2701 GGTCCCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
Db GGTCCCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
QY 2761 CTGCCCCCCCCAAAATGCTTAACCCAGGACCTTGGAATTTCTACTCATCCCAAAATGATAAT 2820  
Db CTGCCCCCCCCAAAATGCTTAACCCAGGACCTTGGAATTTCTACTCATCCCAAAATGATAAT 2820  
QY 2821 TCCAAATGCTGTTTACCAAGTTAGGTGTTGAAGAGGTAGAGGTGGGGTTCAGGT 2880  
Db TCCAAATGCTGTTTACCAAGTTAGGTGTTGAAGAGGTAGAGGTGGGGTTCAGGT 2880  
QY 2881 CTCAACGGCTTCCCTAACCCAGGCTCTCTCTGTTGGCCAGGCTGTTTCCCGGCTTCCA 2940  
Db CTCAACGGCTTCCCTAACCCAGGCTCTCTCTGTTGGCCAGGCTGTTTCCCGGCTTCCA 2940  
QY 2941 CTCCCTCTACTCTCTAGGCTGCTGATGAAGGCACTGCCCAAAATTTCCCGCTTACC 3000  
Db CTCCCTCTACTCTCTAGGCTGCTGATGAAGGCACTGCCCAAAATTTCCCGCTTACC 3000  
QY 3001 CCCAACTTCCCTTACCCCAACTTTCCCAAGCTTCCCAAGCTTCCCAAGCTTCCCAAGCT 3060  
Db CCCAACTTCCCTTACCCCAACTTTCCCAAGCTTCCCAAGCTTCCCAAGCTTCCCAAGCT 3060  
QY 3061 GCAGGACGAGGACAAAGTGGGGTTCCTCAAGGCTTTTCCATCTCAGCCCCCAGAGT 3120  
Db GCAGGACGAGGACAAAGTGGGGTTCCTCAAGGCTTTTCCATCTCAGCCCCCAGAGT 3120  
QY 3121 ATATCTGCTGTTGGGAACTCACAGAACTCAGAGACACCCCTGCTGAGCTAAGG 3180  
Db ATATCTGCTGTTGGGAACTCACAGAACTCAGAGACACCCCTGCTGAGCTAAGG 3180  
QY 3181 GAGGCTTATCTCAGGGGGGTTAAGTGGCTTTGCAATATGCTCTTATTTAT 3240  
Db GAGGCTTATCTCAGGGGGGTTAAGTGGCTTTGCAATATGCTCTTATTTAT 3240  
QY 3241 TAGCGGGTGAATATTTTACTGTAGTGAAGCAATCAGAGTATAATGTTATGTTGAC 3300  
Db TAGCGGGTGAATATTTTACTGTAGTGAAGCAATCAGAGTATAATGTTATGTTGAC 3300







US-09-745-288-100

Query Match 100.0%; Score 3409.6; DB 10; Length 3410;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGACACAGCTGACAGCGCTGGCTCCGGGTGACAGCGCGCTCGGCAGGATCTGA 60  
DB 1 GGGACACAGCTGACAGCGCTGGCTCCGGGTGACAGCGCGCTCGGCAGGATCTGA 60

QY 61 GTGATGAGACGTGTCCCACTAGGTGCCCCACACAGCAGAGTGTGAGCATGGCTGAG 120  
DB 61 GTGATGAGACGTGTCCCACTAGGTGCCCCACACAGCAGAGTGTGAGCATGGCTGAG 120

QY 121 AAGCTGAGACCGGACCAAGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGGAGTT 180  
DB 121 AAGCTGAGACCGGACCAAGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGGAGTT 180

QY 181 GCGGACAGAGGAGGAGCGCGAGCTTCTGGAGCAGAGCGGAGAGGAGGAGTCTG 240  
DB 181 GCGGACAGAGGAGGAGGAGCGCGAGCTTCTGGAGCAGAGCGGAGAGGAGGAGTCTG 240

QY 241 GAGTGCCTGAAGGCGCCCTGAGCCCTACCGCCCTGGCCCACTATGTTCCAGAGCTGTG 300  
DB 241 GAGTGCCTGAAGGCGCCCTGAGCCCTACCGCCCTGGCCCACTATGTTCCAGAGCTGTG 300

QY 301 GGTGAGCGCGCTGCTGGCGCACCGGAAAGCCAGCTCTTGTGTCAACCTGCTAACCTT 360  
DB 301 GGTGAGCGCGCTGCTGGCGCACCGGAAAGCCAGCTCTTGTGTCAACCTGCTAACCTT 360

QY 361 TGGCTGAGAGGTGTGTTGGCGCAGGAGCATCACCTATGTCGCGCTGCTGCTGAGT 420  
DB 361 TGGCTGAGAGGTGTGTTGGCGCAGGAGCATCACCTATGTCGCGCTGCTGCTGAGT 420

QY 421 GGGGTGAGGAGAGTTCATGACCATGTTGCTGGGCATTTGTCAGTGTGCGGCTGGT 480  
DB 421 GGGGTGAGGAGAGTTCATGACCATGTTGCTGGGCATTTGTCAGTGTGCGGCTGGT 480

QY 481 CTGTGTCGCGCTGCTGAGGCTGAGCAGTACCATGTCGCGCTGCTGCTGCTGAGT 540  
DB 481 CTGTGTCGCGCTGCTGAGGCTGAGCAGTACCATGTCGCGCTGCTGCTGCTGAGT 540

QY 541 GCGCTTCATCTGGGCATGTCCTTGGGCATGTCGTCGAGCTCTTTCATCCCAAGGC 600  
DB 541 GCGCTTCATCTGGGCATGTCCTTGGGCATGTCGTCGAGCTCTTTCATCCCAAGGC 600

QY 601 CGGCTGGCTAGCAGGCTGCTGTCGCGGATCCAGGCGCTGAGCTGGCAGTCTCAT 660  
DB 601 CGGCTGGCTAGCAGGCTGCTGTCGCGGATCCAGGCGCTGAGCTGGCAGTCTCAT 660

QY 661 CTGGGCTGGGCTGCTGAGTCTTGTGGCAGGTGCTTCACTCCACTGGAGGCGCT 720  
DB 661 CTGGGCTGGGCTGCTGAGTCTTGTGGCAGGTGCTTCACTCCACTGGAGGCGCT 720

QY 721 GCTCTGACCTTTCGCGGACCCGAGACACTGTGCGCAGGCTACTCTGTCTATGCTT 780  
DB 721 GCTCTGACCTTTCGCGGACCCGAGACACTGTGCGCAGGCTACTCTGTCTATGCTT 780

QY 781 CATGATCAGTCTTGGGGCTGCTGGCTTACCTTCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 781 CATGATCAGTCTTGGGGCTGCTGGCTTACCTTCTGCTGCTGCTGCTGCTGCTGCT 840

QY 841 TGGCTTGGGCTGCTGCTGGGACCCAGGAGGAGTGCCTTCTTGGCTGCTGCTGCTGCT 900  
DB 841 TGGCTTGGGCTGCTGCTGGGACCCAGGAGGAGTGCCTTCTTGGCTGCTGCTGCTGCT 900

QY 901 CTCTCTACCTGCTGCTGAGCAGCAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 901 CTCTCTACCTGCTGCTGAGCAGCAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAG 960

QY 961 CGAGCAGCAGAGGCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 961 CGAGCAGCAGAGGCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

QY 1021 CCCTTGGCTTTCGGAACCTGGCGCCCTGCTTCCCGGCTGCACACAGCTGTGCTGCCG 1080  
DB 1021 CCCTTGGCTTTCGGAACCTGGCGCCCTGCTTCCCGGCTGCACACAGCTGTGCTGCCG 1080

QY 1081 CATGCCCGCACCTCGCGGCTCTTCGTGGCTGAGCTGTGAGCTGTGAGTGTGCTCAT 1140  
DB 1081 CATGCCCGCACCTCGCGGCTCTTCGTGGCTGAGCTGTGAGCTGTGAGTGTGCTCAT 1140

QY 1141 GACCTTACAGCTGTTTACACGATTTGTGGCGGAGGCTGTACAGGCTGTGCCAG 1200  
DB 1141 GACCTTACAGCTGTTTACACGATTTGTGGCGGAGGCTGTACAGGCTGTGCCAG 1200

QY 1201 AGCTGAGCGGCGACCGGCGGAGACACTATGATGAAGGCTTTCGATGGGAGGCT 1260  
DB 1201 AGCTGAGCGGCGACCGGCGGAGACACTATGATGAAGGCTTTCGATGGGAGGCT 1260

QY 1261 GGGCTGTTCTCAGTGGCGCCATCTCCCTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 1320  
DB 1261 GGGCTGTTCTCAGTGGCGCCATCTCCCTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 1320

QY 1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCTGTGCTGC 1380  
DB 1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCTGTGCTGC 1380

QY 1381 CGGTGCCACATGCTGCTGCCACAGTGTGGCGGTGACAGCTTTCAGGCGCCCTCACCGG 1440  
DB 1381 CGGTGCCACATGCTGCTGCCACAGTGTGGCGGTGACAGCTTTCAGGCGCCCTCACCGG 1440

QY 1441 GTTACCTTCTCAGCCTTCAGATTCCTCACACTGGCCTTCCCTTACACCGCGGA 1500  
DB 1441 GTTACCTTCTCAGCCTTCAGATTCCTCACACTGGCCTTCCCTTACACCGCGGA 1500

QY 1501 GAAGCAGGTGCTGCTGCCAAATACCGAGGAGGACCTGGAGGTGCTAGCAGTGTGAGCAG 1560  
DB 1501 GAAGCAGGTGCTGCTGCCAAATACCGAGGAGGACCTGGAGGTGCTAGCAGTGTGAGCAG 1560

QY 1561 CCTGATGACCACTTCTGCGAGGCTTAAAGCCTTGAAGCTTCCCTTCCCTTAAAGCAGCT 1620  
DB 1561 CCTGATGACCACTTCTGCGAGGCTTAAAGCCTTGAAGCTTCCCTTCCCTTAAAGCAGCT 1620

QY 1621 GGGTGTGGAGGAGTGGCTGCTCCACTCCAGCGGCTGTGCGGCGCTGTGCGGCTG 1680  
DB 1621 GGGTGTGGAGGAGTGGCTGCTCCACTCCAGCGGCTGTGCGGCGCTGTGCGGCTG 1680

QY 1681 TGATGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
DB 1681 TGATGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

QY 1741 GGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
DB 1741 GGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800

QY 1801 ATCCCTGTTTATGGCTTCCATTTCCAGCTCAGCAGTGTGCTGCTGCTGCTGCTGCTGCT 1860  
DB 1801 ATCCCTGTTTATGGCTTCCATTTCCAGCTCAGCAGTGTGCTGCTGCTGCTGCTGCTGCT 1860

QY 1861 TGCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
DB 1861 TGCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920

QY 1921 CGACTTGGCCAAATACTCAGCGTAGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
DB 1921 CGACTTGGCCAAATACTCAGCGTAGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980

QY 1981 CACTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
DB 1981 CACTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040

QY 2041 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
DB 2041 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100



Db	241	GAGTGGCTGAAAGCGCCCTGAGCCCTACCCGCGCTGGCCACATATGCTCCAGAGCTGTG	300
Qy	301	GCTGAGCCCTGCTGGGACCCGAAAGCCAGCTCTTGTGTGCTCAACCTGCTTAACCTT	360
Db	301	GGTGAAGCCCTGCTGGGACCCGAAAGCCAGCTCTTGTGTGCTCAACCTGCTTAACCTT	360
Qy	361	TGGGCTGAGAGTGTGTTGGCCGAGGAGATACCTATGATGGCCCTGCTGCTGGAAGT	420
Db	361	TGGGCTGAGAGTGTGTTGGCCGAGGAGATACCTATGATGGCCCTGCTGCTGGAAGT	420
Qy	421	GGGGTACAGAGAGTTCATGACCATGGTCTGGGCAATGGTCCAGTGTGGGCTGGT	480
Db	421	GGGGTACAGAGAGTTCATGACCATGGTCTGGGCAATGGTCCAGTGTGGGCTGGT	480
Qy	481	CTGTGTCCTGCTTAGGCTCAGCCAGTGAACCATGGCGTGGAGCGTATGGCCGCGCGG	540
Db	481	CTGTGTCCTGCTTAGGCTCAGCCAGTGAACCATGGCGTGGAGCGTATGGCCGCGCGG	540
Qy	541	GGCCTTCATCTGGGCACTGCTTGGGCACTCTGCTGAGCCCTTTCTCATCCCAAGGC	600
Db	541	GGCCTTCATCTGGGCACTGCTTGGGCACTCTGCTGAGCCCTTTCTCATCCCAAGGC	600
Qy	601	CGGCTGGCTAGCAGGCTGCTGTGCCCGGATCCAGGCCCTGGAGCTGGCACTGCTCAT	660
Db	601	CGGCTGGCTAGCAGGCTGCTGTGCCCGGATCCAGGCCCTGGAGCTGGCACTGCTCAT	660
Qy	661	CGTGGGCTGGGCTGTGGAGTCTGTGGCCAGAGTGTGCTTCACTGAGGAGGCT	720
Db	661	CGTGGGCTGGGCTGTGGAGTCTGTGGCCAGAGTGTGCTTCACTGAGGAGGCT	720
Qy	721	GCTCTGACCTCTCCGGACCCGAGGAGTGTGCCAGGCTACTGCTGCTATGCTT	780
Db	721	GCTCTGACCTCTCCGGACCCGAGGAGTGTGCCAGGCTACTGCTGCTATGCTT	780
Qy	781	CATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCCATGCACTGGGACACAG	840
Db	781	CATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCCATGCACTGGGACACAG	840
Qy	841	TGCGCTGGCCCTACTTGGGACCCAGGAGGAGTGCCTCTTGGGCTGCTCAACCTCAT	900
Db	841	TGCGCTGGCCCTACTTGGGACCCAGGAGGAGTGCCTCTTGGGCTGCTCAACCTCAT	900
Qy	901	CTTCTCAGCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAGGAGGCTGGGCCAC	960
Db	901	CTTCTCAGCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAGGAGGCTGGGCCAC	960
Qy	961	CGAGCCAGAGAGGCTGTGGGCCCTCTTGTGGCCCTGCTGCTGCTGCTGCTGCTG	1020
Db	961	CGAGCCAGAGAGGCTGTGGGCCCTCTTGTGGCCCTGCTGCTGCTGCTGCTGCTG	1020
Qy	1021	CGGCTGGCTTTCGGGACCTGGGCGCCCTGCTTCCCGGCTGACAGCTGCTGCTGCG	1080
Db	1021	CGGCTGGCTTTCGGGACCTGGGCGCCCTGCTTCCCGGCTGACAGCTGCTGCTGCG	1080
Qy	1081	CATGCCCCGACCTCGCCGCTCTTGTGGCTGAGCTGTGCTGCTGCTGCTGCTGCTG	1140
Db	1081	CATGCCCCGACCTCGCCGCTCTTGTGGCTGAGCTGTGCTGCTGCTGCTGCTGCTG	1140
Qy	1141	GACCTTCAGCTGTTTACAGGATTCGTGGGCGAGGCTGTACCAGGCGTGGCCAG	1200
Db	1141	GACCTTCAGCTGTTTACAGGATTCGTGGGCGAGGCTGTACCAGGCGTGGCCAG	1200
Qy	1201	AGCTGAGCGGGCACCGAGCCCGGAGACACTATGATGAAGGCTTCGATGGGAGCT	1260
Db	1201	AGCTGAGCGGGCACCGAGCCCGGAGACACTATGATGAAGGCTTCGATGGGAGCT	1260
Qy	1261	GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
Db	1261	GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
Qy	1321	GCAGCGATTCGGCACTGAGCACTATTTGGGCGAGTGTGGGAGCTTCCCTGCTGCTG	1380
Db	1321	GCAGCGATTCGGCACTGAGCACTATTTGGGCGAGTGTGGGAGCTTCCCTGCTGCTG	1380
Qy	1381	CGTGGCCATGCTTCCACAGTGTGGCCCTGGTGACAGCTTACGCGCCCTCACCGG	1440
Db	1381	CGTGGCCATGCTTCCACAGTGTGGCCCTGGTGACAGCTTACGCGCCCTCACCGG	1440
Qy	1441	GTTACCTTCTCAGCCCTGAGATCTGCTGCTTACACACTGGGCTCTCTTACACCGGGA	1500
Db	1441	GTTACCTTCTCAGCCCTGAGATCTGCTGCTTACACACTGGGCTCTCTTACACCGGGA	1500
Qy	1501	GAAGCAGGTGTTCTCCGCAATACCGAGGGGACACTGGAGTGTGCTGAGGACAG	1560
Db	1501	GAAGCAGGTGTTCTCCGCAATACCGAGGGGACACTGGAGTGTGCTGAGGACAG	1560
Qy	1561	CTGTATGACAGCTTCTGCGAGGCTTAAGCTGAGCTTCCCTTAAATGGACACCT	1620
Db	1561	CTGTATGACAGCTTCTGCGAGGCTTAAGCTGAGCTTCCCTTAAATGGACACCT	1620
Qy	1621	GGTGTGAGGAGTGGCTTCTCCACCTCCACCCGGGCTGCTGCTGCTGCTGCTGCTG	1680
Db	1621	GGTGTGAGGAGTGGCTTCTCCACCTCCACCCGGGCTGCTGCTGCTGCTGCTGCTG	1680
Qy	1681	TGATGCTCCGTAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1740
Db	1681	TGATGCTCCGTAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1740
Qy	1741	GGCATGCTGCTGACCTGCGCATCTGATGATGCTTCTGCTGCTGCTGCTGCTGCTG	1800
Db	1741	GGCATGCTGCTGACCTGCGCATCTGATGATGCTTCTGCTGCTGCTGCTGCTGCTG	1800
Qy	1801	ATCCTCTTTATGGCTTCCATTTCCAGCTAGCAGTGTGCTGCTGCTGCTGCTGCTG	1860
Db	1801	ATCCTCTTTATGGCTTCCATTTCCAGCTAGCAGTGTGCTGCTGCTGCTGCTGCTG	1860
Qy	1861	TGCGGAGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1920
Db	1861	TGCGGAGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1920
Qy	1921	CGACTTGGCCAAATACCTCAGCTAGAAATTTCCAGCACATTTGGGGTGGAGGCTGCT	1980
Db	1921	CGACTTGGCCAAATACCTCAGCTAGAAATTTCCAGCACATTTGGGGTGGAGGCTGCT	1980
Qy	1981	CAGTGGTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040
Db	1981	CAGTGGTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040
Qy	2041	TTCTGTTGCTGCCAAAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2100
Db	2041	TTCTGTTGCTGCCAAAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2100
Qy	2101	GCTGCACAGCTGGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2160
Db	2101	GCTGCACAGCTGGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2160
Qy	2161	ACTGGAGGCTTCCAAAGGGGTTTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2220
Db	2161	ACTGGAGGCTTCCAAAGGGGTTTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2220
Qy	2221	ATGCACTGGAATGGGGGCTGCTGAGGTGATTTACCCAGGCTCAGGCTTAAACAGCTAGC	2280
Db	2221	ATGCACTGGAATGGGGGCTGCTGAGGTGATTTACCCAGGCTCAGGCTTAAACAGCTAGC	2280
Qy	2281	CTCCTAGTTGAGACACACTAGAGAGGGTTTTTGGGAGCTGAATAACTCAGTCACTG	2340
Db	2281	CTCCTAGTTGAGACACACTAGAGAGGGTTTTTGGGAGCTGAATAACTCAGTCACTG	2340
Qy	2341	GTTTCCCATCTTACAGCCCTTAACTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2400
Db	2341	GTTTCCCATCTTACAGCCCTTAACTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2400
Qy	2401	TTTCTAGGATGAACACTCTCCATGGGATTTGAACATATGACTTATTTAGGGGAGA	2460
Db	2401	TTTCTAGGATGAACACTCTCCATGGGATTTGAACATATGACTTATTTAGGGGAGA	2460









Db 2761 CTGCCCCCAAAATGCTACCCAGACCTTGAAATTTACTCATCCAAATGATAAT 2820  
Qy 2821 TCACAAATGCTGTTACCAAGGTTAGGCTTTGAAGAAAGTAGAGGTGGGCTTCAGGT 2880  
Db 2821 TCACAAATGCTGTTACCAAGGTTAGGCTTTGAAGAAAGTAGAGGTGGGCTTCAGGT 2880  
Qy 2881 CTCACGGCTTCCCTAACCAACCCCTCTCTCTTTGGCCAGCCTGTTCCCCACATCCA 2940  
Db 2881 CTCACGGCTTCCCTAACCAACCCCTCTCTCTTTGGCCAGCCTGTTCCCCACATCCA 2940  
Qy 2941 CTCCTCTCTACTCTCTAGACTGGCTGATGAAGCACTGCCCCCAAAATTTCCCTACC 3000  
Db 2941 CTCCTCTCTACTCTCTAGACTGGCTGATGAAGCACTGCCCCCAAAATTTCCCTACC 3000  
Qy 3001 CCCAACTTTCCCTACCCCAACTTTCCCAACAGCTTCCCAACCCCTGTTGGAGTACT 3060  
Db 3001 CCCAACTTTCCCTACCCCAACTTTCCCAACAGCTTCCCAACCCCTGTTGGAGTACT 3060  
Qy 3061 GCAGGACCAAGCAACAAAGTGGGTTTCCCAAGCCTTTGTCATCTCAGCCCCCAGAGT 3120  
Db 3061 GCAGGACCAAGCAACAAAGTGGGTTTCCCAAGCCTTTGTCATCTCAGCCCCCAGAGT 3120  
Qy 3121 ATATCTGTGCTTGGGGAATCTACACAGAACTCAGAGCAACCCCTGCTGAGCTAAGG 3180  
Db 3121 ATATCTGTGCTTGGGGAATCTACACAGAACTCAGAGCAACCCCTGCTGAGCTAAGG 3180  
Qy 3181 GAGTCTTATCTCAGGGGGGTTTAAAGTGGGTTTGAATGCTGCTTATTTAT 3240  
Db 3181 GAGTCTTATCTCAGGGGGGTTTAAAGTGGGTTTGAATGCTGCTTATTTAT 3240  
Qy 3241 TAGCGGGTGAATATTTTATCTGTAAGTGAGCAATCAGATGATATGTTATGTGACA 3300  
Db 3241 TAGCGGGTGAATATTTTATCTGTAAGTGAGCAATCAGATGATATGTTATGTGACA 3300  
Qy 3301 AAATTAAGGCTTTCTTATATGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3360  
Db 3301 AAATTAAGGCTTTCTTATATGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3360  
Qy 3361 AAAAAAATAA 3410  
Db 3361 AAAAAAATAA 3410

## RESULT 9

US-09-030-606-110  
: Sequence 110, Application US/09030606  
: Patent No. US20020081580A1  
: GENERAL INFORMATION:  
: APPLICANT: Xu, Jiangchun  
: APPLICANT: Dillon, Davin C.  
: TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS F  
: NUMBER OF SEQUENCES: 224  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: SEED AND BERRY LLP  
: STREET: 6300 Columbia Center, 701 Fifth Avenue  
: CITY: Seattle  
: STATE: WA  
: COUNTRY: USA  
: ZIP: 98104  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/030,606  
: FILING DATE: 25-FEB-1998  
: CLASSIFICATION:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Maki, David J.  
: REGISTRATION NUMBER: 31,392  
: REFERENCE/DOCKET NUMBER: 210121.428C3

: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (206) 622-4900  
: TELEFAX: (206) 682-6031  
: INFORMATION FOR SEQ ID NO: 110:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 3410 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: cDNA  
: ORIGINAL SOURCE:  
: ORGANISM: Homo sapiens  
: US-09-030-606-110

Query Match 100.0%; Score 3409.6; DB 10; Length 3410;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGGAAACCAAGCCTGACGCGCTGGTCCGGGTGACAGCCGCGCTCGGCGAGGATCTGA 60  
Db 1 GGGAAACCAAGCCTGACGCGCTGGTCCGGGTGACAGCCGCGCTCGGCGAGGATCTGA 60  
Qy 61 GTGATGACAGCTGTCCTCCACTAGGTGCCCCACACAGCAGAGGTGTTGAGCATGGGCTGAG 120  
Db 61 GTGATGACAGCTGTCCTCCACTAGGTGCCCCACACAGCAGAGGTGTTGAGCATGGGCTGAG 120  
Qy 121 AAGCTGGACCGCCACCAAGGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGCAGTT 180  
Db 121 AAGCTGGACCGCCACCAAGGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGCAGTT 180  
Qy 181 GCGGCGACAAAGGAGGAGGCGCCAGCTTCTGGAGCAGAGCCGAGCAGAGCAAGTCTGT 240  
Db 181 GCGGCGACAAAGGAGGAGGCGCCAGCTTCTGGAGCAGAGCCGAGCAGAGCAAGTCTGT 240  
Qy 241 GAGTGCCTGAAAGCGCCCTGAGCCCTACCGCTGCCCCCTGCCCCACTATGTTCCAGAGGCTGTG 300  
Db 241 GAGTGCCTGAAAGCGCCCTGAGCCCTACCGCTGCCCCCTGCCCCACTATGTTCCAGAGGCTGTG 300  
Qy 301 GGTGAGCGCCTGCTGGGCGACCGAAAGCCAGCTCTTGTGCTCAACCTGCTAAGCTT 360  
Db 301 GGTGAGCGCCTGCTGGGCGACCGAAAGCCAGCTCTTGTGCTCAACCTGCTAAGCTT 360  
Qy 361 TGGCCTGGAGGTGTTTGGCCGACGATCACCTATGTCGCCCTCTGCTGCTGGAAGT 420  
Db 361 TGGCCTGGAGGTGTTTGGCCGACGATCACCTATGTCGCCCTCTGCTGCTGGAAGT 420  
Qy 421 GGGGTAGAGGAGAAAGTTTCATGACCATGTTGGTGGCATTTGGTCCAGTGTGCGCTTGGT 480  
Db 421 GGGGTAGAGGAGAAAGTTTCATGACCATGTTGGTGGCATTTGGTCCAGTGTGCGCTTGGT 480  
Qy 481 CTGTGTCGGCTCTAGGCTCAGCCAGTACCATGCGGTGGAGCTATGGCCGCGCCG 540  
Db 481 CTGTGTCGGCTCTAGGCTCAGCCAGTACCATGCGGTGGAGCTATGGCCGCGCCG 540  
Qy 541 GCCCTTCATCTGGGCACTGCTTGGGCACTGCTGAGCCTCTTCTTCATCCCAAGGC 600  
Db 541 GCCCTTCATCTGGGCACTGCTTGGGCACTGCTGAGCCTCTTCTTCATCCCAAGGC 600  
Qy 601 CGGCTGCTAGCAGGGCTGCTGTCGCCGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
Db 601 CGGCTGCTAGCAGGGCTGCTGTCGCCGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
Qy 661 CTTGGGGCTGGGCTGCTGAGCTTCTGTGGCCAGGTGTTGTTCACTCCACTGAGAGCCCT 720  
Db 661 CTTGGGGCTGGGCTGCTGAGCTTCTGTGGCCAGGTGTTGTTCACTCCACTGAGAGCCCT 720  
Qy 721 GCTCTCTGACCTCTTCCGGGACCCGACCACTGTCGCCAGGCCCTACTCTGCTATGCCCT 780  
Db 721 GCTCTCTGACCTCTTCCGGGACCCGACCACTGTCGCCAGGCCCTACTCTGCTATGCCCT 780  
Qy 781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTCCTGCCCTGCCATTTGACTGGGACACAG 840  
Db 781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTCCTGCCCTGCCATTTGACTGGGACACAG 840





1441 GTTACACCTCTCAGCCCTGAGATCCTGCCCTACACACTGGCCCTCCTTACACCGGGA 1500  
1441 GTTACACCTCTCAGCCCTGAGATCCTGCCCTACACACTGGCCCTCCTTACACCGGGA 1500  
1501 GAAGCAGGTGTTCTGCCCAATACGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560  
1501 GAAGCAGGTGTTCTGCCCAATACGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560  
1561 CTTGATGACACAGCTTCTGTCAGGCGCTTAAGCCTGGAGCTCCCTTAATGGACAGT 1620  
1561 CTTGATGACACAGCTTCTGTCAGGCGCTTAAGCCTGGAGCTCCCTTAATGGACAGT 1620  
1621 GGTGCTGGAGGAGTGGCTGCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCCCTG 1680  
1621 GGTGCTGGAGGAGTGGCTGCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCCCTG 1680  
1681 TGATGCTCCGCTACGTTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
1681 TGATGCTCCGCTACGTTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
1741 GGGATGCTGCTGACCTGCGCATCTGCTGGATAGTGGCTTCTGCTGCTGCTGCTGCTGCTG 1800  
1741 GGGATGCTGCTGACCTGCGCATCTGCTGGATAGTGGCTTCTGCTGCTGCTGCTGCTGCTG 1800  
1801 ATCCCTGTTTATGGGCTCCATTTGTCAGCTCAGCCAGTCTGTCACCTGCTATATGCTGTC 1860  
1801 ATCCCTGTTTATGGGCTCCATTTGTCAGCTCAGCCAGTCTGTCACCTGCTATATGCTGTC 1860  
1861 TGCGCAGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
1861 TGCGCAGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
1921 CGACTTGGCAATACCTACAGCTAGAACTTCCAGCACATTTGGGTGGAGGCTTGCCT 1980  
1921 CGACTTGGCAATACCTACAGCTAGAACTTCCAGCACATTTGGGTGGAGGCTTGCCT 1980  
1981 CACTGGGTCCAGCTCCCGCTCTGCTGTTAGCCCATGGGCTGCTGCTGCTGCTGCTGCTGCT 2040  
1981 CACTGGGTCCAGCTCCCGCTCTGCTGTTAGCCCATGGGCTGCTGCTGCTGCTGCTGCTGCT 2040  
2041 TTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
2041 TTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
2101 GCTGACAGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTG 2160  
2101 GCTGACAGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTG 2160  
2161 ACTGGAGGCTTCCAGGGGCTTTCAGTCTGGACTTATACAGGAGGAGGAGGAGGAGGAGGAG 2220  
2161 ACTGGAGGCTTCCAGGGGCTTTCAGTCTGGACTTATACAGGAGGAGGAGGAGGAGGAGGAG 2220  
2221 ATGACCTGGAATGCGGAGCTCTGACAGGTGATTAACAGGCTGAGGCTGAGGCTGAGGCTGAG 2280  
2221 ATGACCTGGAATGCGGAGCTCTGACAGGTGATTAACAGGCTGAGGCTGAGGCTGAGGCTGAG 2280  
2281 CTCCTAGTTGAGACACACCTAGAGAGGCTTTTGGAGCTGATTAACCTCAGTCACTGCTGCTG 2340  
2281 CTCCTAGTTGAGACACACCTAGAGAGGCTTTTGGAGCTGATTAACCTCAGTCACTGCTGCTG 2340  
2341 GTTTCCTCCTCTTAAGCCCTTAACTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
2341 GTTTCCTCCTCTTAAGCCCTTAACTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
2401 TTTCTAGATGAACACTCTCCATGGGATTTGAACATATGACTTATTTAGGGAAGA 2460  
2401 TTTCTAGATGAACACTCTCCATGGGATTTGAACATATGACTTATTTAGGGAAGA 2460  
2461 GTCCTGAGGGGCAACACACAGACAGGCTCCCTGACCCACAGCACTCTCTTTGCT 2520  
2461 GTCCTGAGGGGCAACACACAGACAGGCTCCCTGACCCACAGCACTCTCTTTGCT 2520

2521 GATCACCCTCTTACCTTTTATCAGATGTCGCTGTTGGTCTCTCTCTCTCTCTCTCTCTCT 2580  
2521 GATCACCCTCTTACCTTTTATCAGATGTCGCTGTTGGTCTCTCTCTCTCTCTCTCTCTCT 2580  
2581 CAGACACAGGCACTTTAAATTTATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTT 2640  
2581 CAGACACAGGCACTTTAAATTTATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTT 2640  
2641 TGCTAGCTTTCTCTGTTGTTGCTCTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2700  
2641 TGCTAGCTTTCTCTGTTGTTGCTCTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2700  
2701 GGTCCCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
2701 GGTCCCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
2761 CTGGCCCCCAAAATGCTTAACCCAGGCTTTGAAAGTGTGAAGGAGGTAGAGGTTGGGCTTTC 2820  
2761 CTGGCCCCCAAAATGCTTAACCCAGGCTTTGAAAGTGTGAAGGAGGTAGAGGTTGGGCTTTC 2820  
2821 TCCAAATGCTGTTACCCAGGCTTTAGGTTGTTGAAAGGAGGTAGAGGTTGGGCTTTCAGGT 2880  
2821 TCCAAATGCTGTTACCCAGGCTTTAGGTTGTTGAAAGGAGGTAGAGGTTGGGCTTTCAGGT 2880  
2881 CTCAACGGCTTCCCTTAACCCAGGCTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940  
2881 CTCAACGGCTTCCCTTAACCCAGGCTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940  
2941 CTCCCTCTACTCTCTTAGGCTGGCTGATGAAGGACTGCCCAAAATTTCCCTTACC 3000  
2941 CTCCCTCTACTCTCTTAGGCTGGCTGATGAAGGACTGCCCAAAATTTCCCTTACC 3000  
3001 CCCAACTTTCCCTTACCCCAACTTTCCCAAGCTTTCCCAAGCTTTCCCAAGCTTTCCCAAGCT 3060  
3001 CCCAACTTTCCCTTACCCCAACTTTCCCAAGCTTTCCCAAGCTTTCCCAAGCTTTCCCAAGCT 3060  
3061 GCAGGACCAAGAGCAAAAGTGGCTTTCCCAAGCTTTCCCAAGCTTTCCCAAGCTTTCCCAAGCT 3120  
3061 GCAGGACCAAGAGCAAAAGTGGCTTTCCCAAGCTTTCCCAAGCTTTCCCAAGCTTTCCCAAGCT 3120  
3121 ATATCTGCTTTGGGGAATCTCACAGAACTCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180  
3121 ATATCTGCTTTGGGGAATCTCACAGAACTCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180  
3181 GAGGTCTTATCTCTCAGGGGGGTTTAACTGCTGCTTAACTGCTGCTTAACTGCTTAACTGCT 3240  
3181 GAGGTCTTATCTCTCAGGGGGGTTTAACTGCTGCTTAACTGCTGCTTAACTGCTTAACTGCT 3240  
3241 TAGCGGGTGAATTTTATATCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3300  
3241 TAGCGGGTGAATTTTATATCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3300  
3301 AAATTAAGGCTTTCTTATATCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3360  
3301 AAATTAAGGCTTTCTTATATCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3360  
3361 AAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAA 3410  
3361 AAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAA 3410

## RESULT 11

US-09-115-453-110  
; Sequence 110, Application US/09115453B  
; Patent No. US20020090372A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillion, Davin C.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND  
; TITLE OF INVENTION: METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.427C4  
; CURRENT APPLICATION NUMBER: US/09/115,453B  
; CURRENT FILING DATE: 1998-07-14



```

1981  CACTGGTCCAGCTCCCGCTCTCTGTTAGCCCATGGGGCTCGCGGCTGGCCGCGAGT 2040
QY      |||
2041  TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCGCACCCCTGTGCTGCTGAGTGGTA 2100
Db      |||
2041  TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCGCACCCCTGTGCTGCTGAGTGGTA 2100
QY      |||
2101  GCTGCACAGCTGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 2160
Db      |||
2101  GCTGCACAGCTGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 2160
QY      |||
2161  ACTGAGGCTTCCAAAGGGGTTTCAGTCTGGACTTATACAGGAGGCGCCAGAGGGCTCC 2220
Db      |||
2161  ACTGAGGCTTCCAAAGGGGTTTCAGTCTGGACTTATACAGGAGGCGCCAGAGGGCTCC 2220
QY      |||
2221  ATGCACTGGAATGCGGGGACTCTGAGTGGATACCCAGGCTCAGGGTTAACAGCTAGC 2280
Db      |||
2221  ATGCACTGGAATGCGGGGACTCTGAGTGGATACCCAGGCTCAGGGTTAACAGCTAGC 2280
QY      |||
2281  CTCCTAGTTGAGACACACTAGAGAGGGTTTGGGAGCTGAATTAACCTCAGTCACCTG 2340
Db      |||
2281  CTCCTAGTTGAGACACACTAGAGAGGGTTTGGGAGCTGAATTAACCTCAGTCACCTG 2340
QY      |||
2341  GTTCCCACTCTAAGCCCTTAACTGCAGCTTCGTTTAATGATAGCTCTTTCATGGGAG 2400
Db      |||
2341  GTTCCCACTCTAAGCCCTTAACTGCAGCTTCGTTTAATGATAGCTCTTTCATGGGAG 2400
QY      |||
2401  TTCTAGGATGAACACTCTCCATGGGATTTGAACATATGACTTATTTGTTAGGGAGA 2460
Db      |||
2401  TTCTAGGATGAACACTCTCCATGGGATTTGAACATATGACTTATTTGTTAGGGAGA 2460
QY      |||
2461  GTCTAGGAGGCAACACACAGAACAGCTCCCTCAGCCACACAGCTGCTTTTTCCT 2520
Db      |||
2461  GTCTAGGAGGCAACACACAGAACAGCTCCCTCAGCCACACAGCTGCTTTTTCCT 2520
QY      |||
2521  GATCCACCCCTCTTACCTTTTATCAGGATGGGCTGTGGTCCCTCTCTGTGCGATCA 2580
Db      |||
2521  GATCCACCCCTCTTACCTTTTATCAGGATGGGCTGTGGTCCCTCTCTGTGCGATCA 2580
QY      |||
2581  CAGACACAGCATTAAATATTAATTAATTTTAACTATTTTAACTAGAGGGAATCCAT 2640
Db      |||
2581  CAGACACAGCATTAAATATTAATTAATTTTAACTATTTTAACTAGAGGGAATCCAT 2640
QY      |||
2641  TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGTGTGGGATGCGGATCCCAACAATCA 2700
Db      |||
2641  TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGTGTGGGATGCGGATCCCAACAATCA 2700
QY      |||
2701  GGTCCCTCAGATAGTGTGCTATGGGCTGATCATTCAGAACTCTTCTCTCTGCGGT 2760
Db      |||
2701  GGTCCCTCAGATAGTGTGCTATGGGCTGATCATTCAGAACTCTTCTCTCTGCGGT 2760
QY      |||
2761  CTGGCCGCCCAAAATGCCTAACCCAGGACCTTGGAAATTTCTATCCCAAAATGATAAT 2820
Db      |||
2761  CTGGCCGCCCAAAATGCCTAACCCAGGACCTTGGAAATTTCTATCCCAAAATGATAAT 2820
QY      |||
2821  TCCAAATGCTGTGTACCAAGGTTAGGTTGTTGAAGGAAGTTAGAGGGTGGGCTTCAGT 2880
Db      |||
2821  TCCAAATGCTGTGTACCAAGGTTAGGTTGTTGAAGGAAGTTAGAGGGTGGGCTTCAGT 2880
QY      |||
2881  CTCAGGGCTTCCCTAACCAACCCCTCTTCTCTGCGGAGCTGCTTCCGCCCACTTCA 2940
Db      |||
2881  CTCAGGGCTTCCCTAACCAACCCCTCTTCTCTGCGGAGCTGCTTCCGCCCACTTCA 2940
QY      |||
2941  CTCCTCTACTCTCTAGACTGGGCTGATCAAGGCACTGCCCAAAATTTCCCTACC 3000
Db      |||
2941  CTCCTCTACTCTCTAGACTGGGCTGATCAAGGCACTGCCCAAAATTTCCCTACC 3000
QY      |||
3001  CCCAACTTCCCTTACCCCAACTTTCCCAACAGCTTCCCAACCCCTGTTTGGAGTACT 3060
Db      |||
3001  CCCAACTTCCCTTACCCCAACTTTCCCAACAGCTTCCCAACCCCTGTTTGGAGTACT 3060
QY      |||
3061  GCAGGACCAAGACAAAGTGGGTTTCCCAAGCTTTCCCAAGCTTCTCATCTCAGCCCCCAGAGT 3120
Db      |||
3061  GCAGGACCAAGACAAAGTGGGTTTCCCAAGCTTTCCCAAGCTTCTCATCTCAGCCCCCAGAGT 3120
QY      |||

RESULT 12
US-09-838-785-1
; Sequence 1, Application US/09838785
; Patent No. US20020009455A1
; GENERAL INFORMATION:
; APPLICANT: Lau, Ted
; APPLICANT: Lin, Rick
; APPLICANT: Parkes, Debbie
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; APPLICANT: Van Heuit, Pam T
; APPLICANT: Wu, John
; TITLE OF INVENTION: DNA Encoding a No. US20020009455A1el PROST 03
; FILE REFERENCE: 51831AUSM1
; CURRENT APPLICATION NUMBER: US/09/838.785
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/200,065
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3320
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (282)..(1943)
US-09-838-785-1

Query Match      96.6%; Score 3292.4; DB 10; Length 3320;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3316; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY      3  GAACACAGCTGCACGCGCTGCGTCCGGGTGACAGCGCGCGCTCGGCCAGGATCTGAGT 62
Db      1  GAACACAGCTGCACGCGCTGCGTCCGGGTGACAGCGCGCGCTCGGCCAGGATCTGAGT 60
QY      63  GATGAGAGCTGTCCCACTAGAGTGCCTCCAGAGAGAGTGTGAGCATGGCTGAGAA 122
Db      61  GATGAGAGCTGTCCCACTAGAGTGCCTCCAGAGAGAGTGTGAGCATGGCTGAGAA 120
QY      123  GCTGACCGGCAACCAAGGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGGCAAGTGG 182
Db      121  GCTGACCGGCAACCAAGGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGGCAAGTGG 180
QY      183  CGGACAGAGAGGAGAGAGCGCCAGCTTCTTGAGCAGAGAGCCGAGACAGAGCTTCTGGA 242
Db      181  CGGACAGAGAGGAGAGAGCGCCAGCTTCTTGAGCAGAGAGCCGAGACAGAGCTTCTGGA 240
QY      243  GTGCTGAACGCGCCCTCTAGCCCTTACCGCGCTGGCCCACTATGTTGCCAGAGGCTGTGG 302
```



Db 241 GTCCCTGAACGGGCCCTGAGCCCTACCCGCTGCGCCACTATGCTCCAGAGGCTGTGGG 300  
QY 303 TGAGCCGCGCTGTGCGGCAACCGAAGCCAGCTCTTGCTGTCAACCTGTAACTTTG 362  
Db 301 TGAGCCGCGCTGTGCGGCAACCGAAGCCAGCTCTTGCTGTCAACCTGTAACTTTG 360  
QY 363 GCTGAGGTGTGTTTGGCGGAGGCAATCACTATGTGCGGCTCTGCTGTGGAAGTGG 422  
Db 361 GCTGAGGTGTGTTTGGCGGAGGCAATCACTATGTGCGGCTCTGCTGTGGAAGTGG 420  
QY 423 GGTAGAGAGAAGTTATGACCAATGGTGTGGGCAATGGTCCAGTGTGGGCTGTGGTCT 482  
Db 421 GGTAGAGAGAAGTTATGACCAATGGTGTGGGCAATGGTCCAGTGTGGGCTGTGGTCT 480  
QY 483 GTGTCCCGCTCTAGGCTCAGCACTGACCACTGCGCTGGAGCTATGGCGCGCCGCGG 542  
Db 481 GTGTCCCGCTCTAGGCTCAGCACTGACCACTGCGCTGGAGCTATGGCGCGCCGCGG 540  
QY 543 CTTTATCTGGGCACTGTCTTGGGCACTCTGCTGAGCCTCTTTCTATCCAAAGGCGG 602  
Db 541 CTTTATCTGGGCACTGTCTTGGGCACTCTGCTGAGCCTCTTTCTATCCAAAGGCGG 600  
QY 603 GCTGCTAGAGGCTGTGTCGCGGATCCAGGCGCTGGAGCTGGCACTGCTCATCC 662  
Db 601 GCTGCTAGAGGCTGTGTCGCGGATCCAGGCGCTGGAGCTGGCACTGCTCATCC 660  
QY 663 TGGGCTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 722  
Db 661 TGGGCTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 720  
QY 723 TCTCTGACCTCTTCCGGGACCGGACCACTGTGCGGAGGCTGTCTATGCCCTCA 782  
Db 721 TCTCTGACCTCTTCCGGGACCGGACCACTGTGCGGAGGCTGTCTATGCCCTCA 780  
QY 783 TCATCAGTCTTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 842  
Db 781 TCATCAGTCTTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 840  
QY 843 CCTTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 902  
Db 841 CCTTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 900  
QY 903 TCTCAGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 962  
Db 901 TCTCAGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 960  
QY 963 AGCCAGCAGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1022  
Db 961 AGCCAGCAGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1020  
QY 1023 GCTTGGCTTTCGGGAACTGGGCGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACT 1082  
Db 1021 GCTTGGCTTTCGGGAACTGGGCGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACT 1080  
QY 1083 TGCCCGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1142  
Db 1081 TGCCCGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1140  
QY 1143 CTTTACAGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1202  
Db 1141 CTTTACAGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1200  
QY 1203 CTGAGCGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1262  
Db 1201 CTGAGCGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1260  
QY 1263 GCTTCTCTGAGTGGCGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1322  
Db 1261 GCTTCTCTGAGTGGCGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1320  
QY 1323 AGCGATTCGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1382  
Db 1321 AGCGATTCGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1380

Db 1321 AGCGATTCGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1380  
QY 1383 GTGCGCATGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1442  
Db 1381 GTGCGCATGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1440  
QY 1443 TCACCTTCTGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1502  
Db 1441 TCACCTTCTGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1500  
QY 1503 AGCAGGTGTTTCCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1562  
Db 1501 AGCAGGTGTTTCCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1560  
QY 1563 TGATGACCACTGCTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1622  
Db 1561 TGATGACCACTGCTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1620  
QY 1623 GTGCTGAGGCACTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1682  
Db 1621 GTGCTGAGGCACTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1680  
QY 1683 ATGCTCTCCGCTGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1742  
Db 1681 ATGCTCTCCGCTGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1740  
QY 1743 GCATCTCCGCTGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1802  
Db 1741 GCATCTCCGCTGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1800  
QY 1803 CCTTGTATGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1862  
Db 1801 CCTTGTATGAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1860  
QY 1863 CCGCAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1922  
Db 1861 CCGCAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1920  
QY 1923 ACTTGGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1982  
Db 1921 ACTTGGGCACTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 1980  
QY 1983 CTGGGCTCCAGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2042  
Db 1981 CTGGGCTCCAGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2040  
QY 2043 CTGTGCTGCAAGTAAATGTGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2102  
Db 2041 CTGTGCTGCAAGTAAATGTGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2100  
QY 2103 TGCAAGCTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2162  
Db 2101 TGCAAGCTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2160  
QY 2163 TGCAAGCTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2222  
Db 2161 TGCAAGCTGGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2220  
QY 2223 GCATGGAATGCGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2282  
Db 2221 GCATGGAATGCGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2280  
QY 2283 CTTAGTGTGACACACCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2342  
Db 2281 CTTAGTGTGACACACCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2340  
QY 2343 TTTCCCATCTTAAAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2402  
Db 2341 TTTCCCATCTTAAAGGCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2400  
QY 2403 TCTAGGATGAACACACCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2460  
Db 2401 TCTAGGATGAACACACCTGTGTCGCGGATCCAGGCTGGAGCTGGCACTGCTCATCC 2460





541	Db	CTTCATCTGGGCACTGCTCTTTGGGAGTCTCTGCTGAGGCTCTTTCTCATGCCAAGGGCCGG	600
604	QY	CTGGCTAGCAGGGTGTGTGCCCGGATCCACAGGCCCTGGAGGTGGCACTGCTCATCCT	663
601	Db	CTGGCTAGCAGGGTGTGTGCCCGGATCCACAGGCCCTGGAGGTGGCACTGCTCATCCT	660
664	QY	GGGCGTGGGGCTGCTGACATTTCTGTGGCCAGGTGTCTCACTCCACTGGAGGCCCTGCT	723
661	Db	GGGCGTGGGGCTGCTGACATTTCTGTGGCCAGGTGTCTCACTCCACTGGAGGCCCTGCT	720
724	QY	CTCTGACCTCTTCCGGGACCCGGACCACTGTGCGCAGGCCCTACTCTCTCATGCTTCAT	783
721	Db	CTCTGACCTCTTCCGGGACCCGGACCACTGTGCGCAGGCCCTACTCTCTCATGCTTCAT	780
784	QY	GATCAGTCTTGGGGCTGCTTGCGCTACCTCTGCTGCCATTGACTGGACACCACTGC	843
781	Db	GATCAGTCTTTGGGGCTGCTTGCGCTACCTCTGCTGCCATTGACTGGACACCACTGC	840
844	QY	CCTGGCCCCCTTACCTGGGCAACCCAGGAGGAGTGCTCTTTGGCTGCTCACCCTCATCTT	903
841	Db	CCTGGCCCCCTTACCTGGGCAACCCAGGAGGAGTGCTCTTTGGCTGCTCACCCTCATCTT	900
904	QY	CCTCACCTGCTAGCACCACTGCTGTGTGGTGTGAGGAGGACGCTGGGCCCCACCGA	963
901	Db	CCTCACCTGCTAGCACCACTGCTGTGTGGTGTGAGGAGGACGCTGGGCCCCACCGA	960
964	QY	GCCAGCAGAAGGCTGTGCGGCCCTCTCTGTGCGCCCACTGCTGTCCATGCGGGGCCG	1023
961	Db	GCCAGCAGAAGGCTGTGCGGCCCTCTCTGTGCGCCCACTGCTGTCCATGCGGGGCCG	1020
1024	QY	CTTGGCTTTCCGGAACCTGGCGGCCCTGCTTCCCGGCTGCACAGCTGTGTCGCGCAT	1083
1021	Db	CTTGGCTTTCCGGAACCTGGCGGCCCTGCTTCCCGGCTGCACAGCTGTGTCGCGCAT	1080
1084	QY	GCCTCGACCTTGGCGCGCTCTTCTGTGGCTGAGCTGTGCACTGGATGGCACTCATGAC	1143
1081	Db	GCCTCGACCTTGGCGCGCTCTTCTGTGGCTGAGCTGTGCACTGGATGGCACTCATGAC	1140
1144	QY	CTTCACGCTGTTTACACGGATTTCGTGGCGAGGGCTGTACACAGGCTGTGCCAGAGC	1203
1141	Db	CTTCACGCTGTTTACACGGATTTCGTGGCGAGGGCTGTACACAGGCTGTGCCAGAGC	1200
1204	QY	TGAGCCGGGACCGAGGCCCGGACACTATGATGA-----	1239
1201	Db	TGAGCCGGGACCGAGGCCCGGACACTATGATGAAGGTAAAGCTTGGCAGCCAGCAG	1260
1240	QY	-----	1239
1261	Db	AGGCTGGTGTGGGACCGCCACCAGACAGCACATCGGGGCTGTCTGGGGCTGGTGCC	1320
1240	QY	-----	1239
1321	Db	TCCTCATCTTGGCCCGGACTTCTCTGTAGGAAAGTGGGATGGACCCCATCTGCATACA	1380
1240	QY	-----	1239
1381	Db	CGGCTTCTATGGGTGTGGAACATCTCTGCTGCGGTTCAGGAAGGCTCTGGCTGCTC	1440
1240	QY	-----	1239
1441	Db	TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAGGAAAGCGGAGCTTATTCAA	1500
1240	QY	-----	1239
1501	Db	AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTTCAGATCTGCCTGGTTCAGCGCCAG	1560
1240	QY	-----	1239
1561	Db	TGTGCCCTCTGCTCCCCAACGACTTTCCAAATAATCTCACACGCGCCTTCAGCTCAGG	1620
1240	QY	-----	1239
1621	Db	CGTCTTAGAGGCTCTTAGAGCCTATGGCAGCTGCTTTGTGTTGCTCCCTCTCACCGCCT	1680

Qy	1240	-----	1239
Db	1681	GTCCTCACAGCTGAGACTCCAGGAAACCTTCAGACTACCTTCCTCTGCTTCAGCAAGG	1740
Qy	1240	-----	1239
Db	1741	GGCCTTCCCAACATCTCTGAGGGTCAGTGGAAACCTAGACTCCCATGCTAGAGTA	1800
Qy	1240	-----	1239
Db	1801	GAAGGGAAGGTGCTGGGACGAGGGCTGTCACAGCAGGTCTCGTGACACGAGGTAC	1860
Qy	1240	-----	1239
Db	1861	CTGTGGTTCGGCCTCTCTATCTCCCTGAGACTGCTCCGACCCCTCCCTCCGAGGCTCTGT	1920
Qy	1240	-----AGCGTTTCGGATGGCAGCCTGGGGCTTCTCTGTCGACT	1277
Db	1921	CTGATGCCCCCTCTCCCTCTGACAGGCTTCGGATGGCAGCCTGGGGCTTCTCTGTCGACT	1980
Qy	1278	GGGCCATCTCCCTGGGCTTCTCTCTGTCATGACCCGGCTGGTGACGAGATTGCGACATC	1337
Db	1981	GGGCCATCTCCCTGGCTCTCTCTGTCATGACCCGGCTGGTGACAGCATTCGGCACATC	2040
Qy	1338	GACAGTCTATTGGCCAGTGGCAGCTTCCCTGTGGCTGCCGTGCCATGCCCTGT	1397
Db	2041	GACAGTCTATTGGCCAGTGGCAGCTTCCCTGTGGCTGCCGTGCCATGCCCTGT	2100
Qy	1398	CCCACAGTGGCCGTGGTGACAGCTTCAGCGGCCCTCACCGGTTTACCTTCTCAGCCC	1457
Db	2101	CCCACAGTGGCCGTGGTGACAGCTTCAGCGGCCCTCACCGGTTTACCTTCTCAGCCC	2160
Qy	1458	TGCAGATCTGCCCTACACACTGGCCCTCCCTTACCACCGGAGAACGAGTGTTCCTGC	1517
Db	2161	TGCAGATCTGCCCTACACACTGGCCCTCCCTTACCACCGGAGAACGAGTGTTCCTGC	2220
Qy	1518	CCAAATACCGAGGGACACTGGAGTGTCTAGCAGTGAGGACAGCCTGATGACCACTCC	1577
Db	2221	CCAAATACCGAGGGACACTGGAGTGTCTAGCAGTGAGGACAGCCTGATGACCACTCC	2280
Qy	1578	TGCAGGCCCTAAGCCTGAGCTCCCTTCCCTAATGGACACGTGGTGTGGAGGAGTG	1637
Db	2281	TGCAGGCCCTAAGCCTGAGCTCCCTTCCCTAATGGACACGTGGTGTGGAGGAGTG	2340
Qy	1638	GCCTGTCCACACTCCACCGCGCTCTGCGGGCCTCTGCCTGTGATGTCCGTCAGTG	1697
Db	2341	GCCTGTCCACACTCCACCGCGCTCTGCGGGCCTCTGCCTGTGATGTCCGTCAGTG	2400
Qy	1698	TGCTGTGGGTGAGCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCCTGGACC	1757
Db	2401	TGCTGTGGGTGAGCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCCTGGACC	2460
Qy	1758	TCGCCATCTTGGATAGTGCTTCTCTGCTGCCAGGTGCCCCATCCCTGTTTATGGCT	1817
Db	2461	TCGCCATCTTGGATAGTGCTTCTCTGCTGCCAGGTGCCCCATCCCTGTTTATGGCT	2520
Qy	1818	CCATGTCCAGCTCAGCCAGTCTGTCACTGCCCTATATGCTGTCTGCCGACGCCCTGGGTC	1877
Db	2521	CCATGTCCAGCTCAGCCAGTCTGTCACTGCCCTATATGCTGTCTGCCGACGCCCTGGGTC	2580
Qy	1878	TGCTCGCCATTTACTTTGCTACACAGGTAGTATTTGACNAGAGCAGCTTGGCCAAATACT	1937
Db	2581	TGCTCGCCATTTACTTTGCTACACAGGTAGTATTTGACNAGAGCAGCTTGGCCAAATACT	2640
Qy	1938	CAGCGTAGAAAACCTCCAGCATTGGGGTGGAGGGCTTGCCTCACTGGGTCCCAAGTCC	1997
Db	2641	CAGCGTAGAAAACCTCCAGCATTGGGGTGGAGGGCTTGCCTCACTGGGTCCCAAGTCC	2700
Qy	1998	CCGCTCCTGTTAGCCCATGGGGCTGCGGGCTGGCGGCTGCTGTCTGCTGCCAAG	2057
Db	2701	CCGCTCCTGTTAGCCCATGGGGCTGCGGGCTGGCGGCTGCTGTCTGCTGCCAAG	2760



181 GGCACAGAGGAGAGAGCCGAGCTTCTGAGCAGAGCCGAGACGAGAGTCTTGAG 240  
244 TGCTGAACGGCCCTGAGCCCTACCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303  
241 TGCTGAACGGCCCTGAGCCCTACCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300  
304 GAGCGCCTGTGGGACCGGAAAGCCAGCTCTTGCTGTCAACCTGTAACTTTGG 363  
301 GAGCGCCTGTGGGACCGGAAAGCCAGCTCTTGCTGTCAACCTGTAACTTTGG 360  
364 CTTGAGGTGTGTTGGCCGAGCATCACCATTATGTGCCGCTCTGCTGTGGAAGTGG 423  
361 CTTGAGGTGTGTTGGCCGAGCATCACCATTATGTGCCGCTCTGCTGTGGAAGTGG 420  
424 GGTAGAGAGAGTTCATGACCATGGTCTGGGCATTGGTCCAGCTGTGGCCCTGGTCTG 483  
421 GGTAGAGAGAGTTCATGACCATGGTCTGGGCATTGGTCCAGCTGTGGCCCTGGTCTG 480  
484 TGTCCCGCTCTAGGCTCAGCCAGTACCATGGCGGTGGAGGCTATGGCCGCCGCCGCC 543  
481 TGTCCCGCTCTAGGCTCAGCCAGTACCATGGCGGTGGAGGCTATGGCCGCCGCCGCC 540  
544 CTTCACTGGGCACTGTCTTGGCATCTCTGCTGAGCCTCTTCTCATCCCAAGGCCGG 603  
541 CTTCACTGGGCACTGTCTTGGCATCTCTGCTGAGCCTCTTCTCATCCCAAGGCCGG 600  
604 CTGGCTAGCAGGGTGTGTGCCGGATCCAGGCCCTGGAGCTGGCACTGTCTCATCT 663  
601 CTGGCTAGCAGGGTGTGTGCCGGATCCAGGCCCTGGAGCTGGCACTGTCTCATCT 660  
664 GGGGTGGGCTGTGGACTTCTGTGCCAGGTGTCTTACATCCACTGGAGGCCCTGCT 723  
661 GGGGTGGGCTGTGGACTTCTGTGCCAGGTGTCTTACATCCACTGGAGGCCCTGCT 720  
724 CTCTGACCTTTCGGGACCCGAGGAGGAGTCTTGGCCCTGCTGCTATGCTTCAT 783  
721 CTCTGACCTTTCGGGACCCGAGGAGGAGTCTTGGCCCTGCTGCTATGCTTCAT 780  
784 GATCAGTCTTGGGGCTGTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 843  
781 GATCAGTCTTGGGGCTGTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
844 CTGGCCCTTACCTGGGACCCGAGGAGGAGTCTTGGCCCTGCTGCTGCTGCTGCT 903  
841 CTGGCCCTTACCTGGGACCCGAGGAGGAGTCTTGGCCCTGCTGCTGCTGCTGCT 900  
904 CCTCAGCTGTAGCAGCCACTGCTGCTGGCTGAGGAGGAGGAGGAGGAGGAGGAG 963  
901 CCTCAGCTGTAGCAGCCACTGCTGCTGGCTGAGGAGGAGGAGGAGGAGGAGGAG 960  
964 CCCAGCAGAGGCTGTGGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023  
961 CCCAGCAGAGGCTGTGGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
1024 CTTGGCTTTCGGAACTGGGCGCTGCTTCCCGGCTGCACAGCTGTGCTGCTGCT 1083  
1021 CTTGGCTTTCGGAACTGGGCGCTGCTTCCCGGCTGCACAGCTGTGCTGCTGCT 1080  
1084 GCCCGCACCTGCGCGGCTTCTGCTGGCTGAGCTGTGAGCTGTGAGCTGTGAGCT 1143  
1081 GCCCGCACCTGCGCGGCTTCTGCTGGCTGAGCTGTGAGCTGTGAGCTGTGAGCT 1140  
1144 CTTACAGCTGTTTACAGGATTCGTGGGAGGAGGCTGTACAGGCGCTGCCAGAGC 1203  
1141 CTTACAGCTGTTTACAGGATTCGTGGGAGGAGGCTGTACAGGCGCTGCCAGAGC 1200  
1204 TGAGCGGGACCGAGGCCGAGACACTATGATGA ----- 1239  
1201 TGAGCGGGACCGAGGCCGAGACACTATGATGAAGTAAAGGCTTGGCAGCCAGCAG 1260  
1240 ----- 1239

1261 AGGCTGGTGTGGAGCGCCACCAGAGAGACACTCGGGGCTGTCTGCGGCTGTGCC 1320  
1240 ----- 1239  
1321 TCTCCATCTCGGCCCGACTTCTCTGTAGGAAAGTGGGATGGACCCCATCTGCATACA 1380  
1240 ----- 1239  
1381 CGGCTTCTCATGGTGTGGAACATCTCTGCTTGGCTTTTCAGSAGGCTCTGGCTGCTC 1440  
1240 ----- 1239  
1441 TAGGAGTCTCATCAGAGTCTTGGCCCGACTTTGACAGAAAGAGCGAGCTTATTCAA 1500  
1240 ----- 1239  
1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTTCAGATCTGCCCTGGTTCAGCGCGAG 1560  
1240 ----- 1239  
1561 TGTGCCCTCTGCTGCCCAACGACTTTTCCAAATAATCTCACCAGCGCTTCCAGCTCAG 1620  
1240 ----- 1239  
1621 CGTCTAGAGAGGCTCTTGAAGCCTATGGCCAGCTGTCTTGTGTCTTCCCTCTCACCGGCT 1680  
1240 ----- 1239  
1681 GTCTTACAGCTGAGACTCCAGGAAACCTTTCAGACTACCTTCCCTCTGCTTTCAGCAAG 1740  
1240 ----- 1239  
1741 GCGTTCGCCACATCTCTGAGGGTCACTGGAAGAACTTAGACTCCCATTTGCTAGAGGTA 1800  
1240 ----- 1239  
1801 GAAAGGGAAGGCTGTGGGAGCAGGGCTGGTCCACAGCAGGCTCTGCTGACAGGATAC 1860  
1240 ----- 1239  
1861 CTGTGTTCCGCTTCTCATCTCCTGAGACTGCTCCGACCTTCCCTCCAGGCTCTGT 1920  
1240 ----- AGGCTTCGATGGCAGCTGGGCTGTTCCTTCGCT 1277  
1921 CTGATGGCCCTCTCCTCTGAGGGCTTGGATGGCAGCTGGGCTGTTCCTTCGCT 1980  
1278 GGGCACTCTCCTGCTTCTCTGCTCATGAGCAGGCTGGTGCAGGATTCGGCAGCTC 1337  
1981 GGGCACTCTCCTGCTTCTCTGCTCATGAGCAGGCTGGTGCAGGATTCGGCAGCTC 2040  
1338 GAGCAGCTATTTGGCCAGTGTGGAGCTTCCCTGTGGCTGCCGGTGCCACATGCTGT 1397  
2041 GAGCAGCTATTTGGCCAGTGTGGAGCTTCCCTGTGGCTGCCGGTGCCACATGCTGT 2100  
1398 CCCAGCTGTGGCGTGGAGCTTTCAGCGCCCTTACCGGGTTCACCTTCTCAGCCC 1457  
2101 CCCAGCTGTGGCGTGGAGCTTTCAGCGCCCTTACCGGGTTCACCTTCTCAGCCC 2160  
1458 TGCAGATCTGCCCTACACACTGGCTCCTCTACACCGGAGAGAGGCTGTTCCTGTC 1517  
2161 TGCAGATCTGCCCTACACACTGGCTCCTCTACACCGGAGAGAGGCTGTTCCTGTC 2220  
1518 CCAATACGAGGGGAGCTGGAGTGTGAGCAGTGGAGACAGCTGTATGACAGCTTCC 1577  
2221 CCAATACGAGGGGAGCTGGAGTGTGAGCAGTGGAGACAGCTGTATGACAGCTTCC 2280  
1578 TCCAGGCTTAAAGCTGGAGCTTCCCTTAAAGTGGAGACAGCTGTGAGGAGCTG 1637  
2281 TCCAGGCTTAAAGCTGGAGCTTCCCTTAAAGTGGAGACAGCTGTGAGGAGCTG 2340  
1638 GCCTGCTCCACTCCACCGGCTCTGGGGGCTCTGCTGCTGTGATGCTCCGCTACG 1697  
2341 GCCTGCTCCACTCCACCGGCTCTGGGGGCTCTGCTGCTGTGATGCTCCGCTACG 2400



```
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 704
; LENGTH: 4034
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-704

Query Match      75.8%; Score 2585.4; DB 9; Length 4034;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 3325; Conservative 0; Mismatches 1; Indels 709; Gaps 3;

QY 4 AACACGCTGACGCGCTGGCTCCGGGTGACAGCGCGCGCTCGGCAGGATCTGAGTG 63
DB 1 AACACGCTGACGCGCTGGCTCCGGGTGACAGCGCGCGCTCGGCAGGATCTGAGTG 60
QY 64 ATGAGACGTGTCCCACTGAGTGCGCCACAGCAGCAGGTGTGAGCATGGCTGAGAAG 123
DB 61 ATGAGACGTGTCCCACTGAGTGCGCCACAGCAGCAGGTGTGAGCATGGCTGAGAAG 120
QY 124 CTGGACCGGACCAAGGCGTGGCAGAAATGGCGCCTGGCTGATTCCTAGGCAAGTTGGC 183
DB 121 CTGGACCGGACCAAGGCGTGGCAGAAATGGCGCCTGGCTGATTCCTAGGCAAGTTGGC 180
QY 184 GGCACGAGGAGGAGCGCGCAGCTTCTGGAGCAGAGCCGAGCAGGACGTCTTGAG 243
DB 181 GGCACGAGGAGGAGCGCGCAGCTTCTGGAGCAGAGCCGAGCAGGACGTCTTGAG 240
QY 244 TGCCTGAACGGCCCTGAGCCCTACCGCCCTGGCCCACTATGTGTCAGAGGCTGTGGGT 303
DB 241 TGCCTGAACGGCCCTGAGCCCTACCGCCCTGGCCCACTATGTGTCAGAGGCTGTGGGT 300
QY 304 GAGCGCCTGCTGGCGGACCGGAAAGCCAGCTCTTGTGTCACCTGCTAACCTTTGG 363
DB 301 GAGCGCCTGCTGGCGGACCGGAAAGCCAGCTCTTGTGTCACCTGCTAACCTTTGG 360
QY 364 CTTGAGGTGTGTTGGCGCGCAGGACATCACCTATGTGCGCCCTGCTGCTGGAAGTGG 423
DB 361 CTTGAGGTGTGTTGGCGCGCAGGACATCACCTATGTGCGCCCTGCTGCTGGAAGTGG 420
QY 424 GGTAGAGAGAGTTTCATGACCATGGTGTGGGCAATGGTCCAGTGTGGGCTGTGCTG 483
DB 421 GGTAGAGAGAGTTTCATGACCATGGTGTGGGCAATGGTCCAGTGTGGGCTGTGCTG 480
QY 484 TGTCCCGCTCTAGCTAGCAGCAGTACCATGCGCTGGGCTGAGGCTATGGCGCCCGCC 543
DB 481 TGTCCCGCTCTAGCTAGCAGCAGTACCATGCGCTGGGCTGAGGCTATGGCGCCCGCC 540
QY 544 CTTATCTGGGCACTGTCTTGGGCACTCCTGCTGAGCCTCTTCTCATCCCAAGGGCGG 603
DB 541 CTTATCTGGGCACTGTCTTGGGCACTCCTGCTGAGCCTCTTCTCATCCCAAGGGCGG 600
QY 604 CTGGCTACAGGCTGTGTCGCGGATCCAGGCGCCTGGAGCTGSCACTGTCTATCCT 663
DB 601 CTGGCTACAGGCTGTGTCGCGGATCCAGGCGCCTGGAGCTGSCACTGTCTATCCT 660
QY 664 GGGCTGGGCTGTGGACTTCTGTGCGCAGGTGCTTCACTCCACTGGAGGCGCTGCT 723
DB 661 GGGCTGGGCTGTGGACTTCTGTGCGCAGGTGCTTCACTCCACTGGAGGCGCTGCT 720
QY 724 CTTGACCTCTTCCGGGACCGGACCACTGTGCGCAGGCGCTACTCTGCTATGCTTCAT 783
DB 721 CTTGACCTCTTCCGGGACCGGACCACTGTGCGCAGGCGCTACTCTGCTATGCTTCAT 780
QY 784 GATCAGCTTGGGGCTGCTGGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843
DB 781 GATCAGCTTGGGGCTGCTGGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 844 CCTGCCCGCTTACCTGGGACCGGAGGAGTGGCTCTTTGGGCTGCTGCTGCTGCTGCT 903
DB 841 CCTGCCCGCTTACCTGGGACCGGAGGAGTGGCTCTTTGGGCTGCTGCTGCTGCTGCT 900
QY 904 CCTCACCTGGGTAGCAGCCACACTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 963
```



Db	1981	GC	CCATCTCCCTGGTCTTCTCTCTCTGTCATGACCGGCTGGTCAGAGCAATTCGCGACTC	2040
Qy	1338	GAGCAGTCTATTTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGTGCCACATCCCTGT	1397	
Db	2041	GAGCAGTCTATTTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGTGCCACATCCCTGT	2100	
Qy	1398	CCCACAGTGTGGCCGTGTGTGACAGCTTCAGCCGCCCTCACCCGGGTTCACCTTCTCAGGCC	1457	
Db	2101	CCCACAGTGTGGCCGTGTGTGACAGCTTCAGCCGCCCTCACCCGGGTTCACCTTCTCAGGCC	2160	
Qy	1458	TGCAGATCTGCCCTACACACTGGCTCCCTCTACACCGGGAGAGCAGGTGTTCTCTGC	1517	
Db	2161	TGCAGATCTGCCCTACACACTGGCTCCCTCTACACCGGGAGAGCAGGTGTTCTCTGC	2220	
Qy	1518	CCAAATACCCAGGGGACACTGGAGTGTGTAGCAGTGAAGACAGCCTGTATGACACGCTTCC	1577	
Db	2221	CCAAATACCCAGGGGACACTGGAGTGTGTAGCAGTGAAGACAGCCTGTATGACACGCTTCC	2280	
Qy	1578	TGCCAGGCCCTTAAGCCTTGAGCTCCCTTCCCTTAATGGACACGTGGTGTGTGAGGCGAGTG	1633	
Db	2281	TGCCAGGCCCTTAAGCCTTGAGCTCCCTTCCCTTAATGGACACGTGGTGTGTGAGGCGAGTG	2340	
Qy	1638	GCCTGCTCCCACTCCACCCGGCTCTCGGGGGCTCTGCCGTGTATGTCTCCGTACGTTG	1697	
Db	2341	GCCTGCTCCCACTCCACCCGGCTCTCGGGGGCTCTGCCGTGTATGTCTCCGTACGTTG	2400	
Qy	1698	TGTTGTTGGTGTAGCCACCGCAGGCCAGGTTGTTCCGGCCGGGGCATCTGCCTGGACC	1757	
Db	2401	TGTTGTTGGTGTAGCCACCGCAGGCCAGGTTGTTCCGGCCGGGGCATCTGCCTGGACC	2460	
Qy	1758	TCGCCATCTGGATAGTGTCTTCCCTGCTGTCCAGGTGGCCCAATCCCTGTTTATGGGCT	1817	
Db	2461	TCGCCATCTGGATAGTGTCTTCCCTGCTGTCCAGGTGGCCCAATCCCTGTTTATGGGCT	2520	
Qy	1818	CCATTTCCAGCTCAGCCAGTCTGTACATGCCTATATGTTGCTGCCGACAGCCCTGGGTC	1877	
Db	2521	CCATTTGTCAGCTCAGCCAGTCTGTACATGCCTATATGTTGCTGCCGACAGCCCTGGGTC	2580	
Qy	1878	TGTCGCGCAATTTACTTTTGTGTACACAGTGTATTTTGACAAGAGCGACTTGGCCAAATACT	1937	
Db	2581	TGTCGCGCAATTTACTTTTGTGTACACAGTGTATTTTGACAAGAGCGACTTGGCCAAATACT	2640	
Qy	1938	CAGCGTAGAAAACTTCCAGCACATTTGGGTGTGAGGCTTGCCTCTACTGGGTGCCAGCTCC	1997	
Db	2641	CAGCGTAGAAAACTTCCAGCACATTTGGGTGTGAGGCTTGCCTCTACTGGGTGCCAGCTCC	2700	
Qy	1998	CCGCTCCTGTTAGCCCATAGGGCTCCGGGGCTGGCCGCCAGTTCTGTTGCTGCCAAAG	2057	
Db	2701	CCGCTCCTGTTAGCCCATAGGGCTCCGGGGCTGGCCGCCAGTTCTGTTGCTGCCAAAG	2760	
Qy	2058	TAATGTGGCTCTGTGTCGCCACCTGTGCTGCTGTAGGTGCGTAGCTGCACAGCTGGGGGC	2117	
Db	2761	TAATGTGGCTCTGTGTCGCCACCTGTGCTGCTGTAGGTGCGTAGCTGCACAGCTGGGGGC	2820	
Qy	2118	TGGGGCTCCCTCTCTCTCTCCCACTCTTCTAGGGCTGCTTACGTGGAGGCTTCTCAAG	2177	
Db	2821	TGGGGCTCCCTCTCTCTCTCCCACTCTTCTAGGGCTGCTTACGTGGAGGCTTCTCAAG	2880	
Qy	2178	GGGTTTTCAGTCTGGACTTATACAGGAGGCCACAGAGGCTCCATGCACGTGAATCGGG	2237	
Db	2881	GGGTTTTCAGTCTGGACTTATACAGGAGGCCACAGAGGCTCCATGCACGTGAATCGGG	2940	
Qy	2238	GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAAACAGTACGCTTCTAGTGTGAGACACA	2297	
Db	2941	GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGTACGCTTCTAGTGTGAGACACA	3000	
Qy	2298	CCTAGAGAAGGGTTTTTTGGGAGCTGAATAAACAACAGTCACCTGGTTTCCCATCTCTAAGC	2357	
Db	3001	CCTAGAGAAGGGTTTTTTGGGAGCTGAATAAACAACAGTCACCTGGTTTCCCATCTCTAAGC	3060	
Qy	2358	CCCTTAACCTGCAGCTTCTGTTTAAATGTAGCTTCTTCATGGGAGTTTCTAGGATGAACAC	2417	
Db	3061	CCCTTAACCTGCAGCTTCTGTTTAAATGTAGCTTCTTCATGGGAGTTTCTAGGATGAACAC	3120	

Qy	2418	TCCTC	ATGGGATTTGAACATATG--ACATTATTTGTAGGGGAAGAGTCTCTGAGGGGCAC	2475
Db	3121	TCCTC	ATGGGATTTGAACATATG--ACATTATTTGTAGGGGAAGAGTCTCTGAGGGGCAC	3180
Qy	2476	ACACA	GAACAGGTCCTCCAGCCACAGCACTGCTTTTGTCTGATCCACCCCTCTCT	2535
Db	3181	ACACA	GAACAGGTCCTCCAGCCACAGCACTGCTTTTGTCTGATCCACCCCTCTCT	3240
Qy	2536	TACCT	TTTATCAGGATGTGGCCCTGTGGTCTTCTGTTGCCATCACAGAGACACAGGCAT	2595
Db	3241	TACCT	TTTATCAGGATGT--GCCCTGTGGTCTTCTGTTGCCATCACAGAGACACAGGCAT	3299
Qy	2596	TTAAAT	ATTTAACTTATTTTAAACAAAGTAAAGGGGAATCCATTGTCTAGCTTTTCTCT	2655
Db	3300	TTAAAT	ATTTAACTTATTTTAAACAAAGTAAAGGGGAATCCATTGTCTAGCTTTTCTCT	3359
Qy	2656	GTTCG	TCTCTAAATTTGGGTAGGGTGGGGATCCCCAAACAATCAGGTCCCTTGAGATAG	2715
Db	3360	GTTCG	TCTCTAAATTTGGGTAGGGTGGGGATCCCCAAACAATCAGGTCCCTTGAGATAG	3419
Qy	2716	CTGTC	ATTTGGGCTGATTCATTCGCAGAAATCTTCTTCTCTGGGGTCTGGCCCCCAAAAT	2775
Db	3420	CTGTC	ATTTGGGCTGATTCATTCGCAGAAATCTTCTTCTCTGGGGTCTGGCCCCCAAAAT	3479
Qy	2776	GCCTA	ACCCAGGACCTTGGAAATTTCTACTCATCCCAAAATGATAAATTCCAAATGCTGTTAC	2835
Db	3480	GCCTA	ACCCAGGACCTTGGAAATTTCTACTCATCCCAAAATGATAAATTCCAAATGCTGTTAC	3539
Qy	2836	CCAAAG	TTAGGGTCTTCAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT	2895
Db	3540	CCAAAG	TTAGGGTCTTCAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT	3599
Qy	2896	AACCA	CCCCCTCTTCTTTGGCCCCAGCCTTGTTCCTCCCACTTCCACTTCCCTCTACTCTC	2955
Db	3600	AACCA	CCCCCTCTTCTTTGGCCCCAGCCTTGTTCCTCCCACTTCCACTTCCCTCTACTCTC	3659
Qy	2956	TCTAG	GACTGGGCTGATGAAGGCACTGCCAAAATTTCCCTCACTCCCTTCCCTCA	3015
Db	3660	TCTAG	GACTGGGCTGATGAAGGCACTGCCAAAATTTCCCTCACTCCCTTCCCTCA	3719
Qy	3016	CCCCA	CTTTCCCCACAGCTCCCAACACCTGTTTGGAGCTACTCGAGGACCAGAGA	3075
Db	3720	CCCCA	CTTTCCCCACAGCTCCCAACACCTGTTTGGAGCTACTCGAGGACCAGAGA	3779
Qy	3076	CAAA	GTCGGTTCCTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTGGG	3135
Db	3780	CAAA	GTCGGTTCCTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTGGG	3839
Qy	3136	GAAT	CTCACAGAAATCAGGAGCACCCCTGCTGAGCTTAAGGGAGGTCCTATCTCTC	3195
Db	3840	GAAT	CTCACAGAAATCAGGAGCACCCCTGCTGAGCTTAAGGGAGGTCCTATCTCTC	3899
Qy	3196	AGGGG	GGGTTTAAAGTGGCGTTTGCATAAATCTGCTTATTTATTTAGCGGGGTGAATAT	3255
Db	3900	AGGGG	GGGTTTAAAGTGGCGTTTGCATAAATCTGCTTATTTATTTAGCGGGGTGAATAT	3959
Qy	3256	TTTAT	ACTGTAAGTGAGCAATCAGAGTATAATGTTTATGTTGACAAAATTAAGGCTTTC	3315
Db	3960	TTTAT	ACTGTAAGTGAGCAATCAGAGTATAATGTTTATGTTGACAAAATTAAGGCTTTC	4019
Qy	3316	TTATAT	GTGTTAAAAA 3330	
Db	4020	TTATAT	GTGTTAAAAA 4034	

RESULT 16  
US-09-759-143-704  
; Sequence 704, Application US/09759143  
; Patent No. US200202248A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.





Db	1681	GTCTCACAGCTGAGACTCCAGGAAACCTTCAGACTACCTTCTCTGCTTCCAGCAAGG	1740
Qy	1240	-----	1239
Db	1741	GGCGTTGCCACATCTCTGAGGGTCAGTGGAGAACTAGACTCCCATTTGTAGAGTA	1800
Qy	1240	-----	1239
Db	1801	GAAGGGGAGGGTGTGGGAGCAGGGCTGGTCCACACAGGCTCTCGTGCAGCAGGTAC	1860
Qy	1240	-----	1239
Db	1861	CTGTGGTTCGGCTTCTCATCTCCCTGAGACTGCTCCGACCTTCCCTCCAGGCTCTGT	1920
Qy	1240	-----	1239
Db	1921	CTGATGGCCCTCTCCCTCTGACAGCGCTTCGATGGGAGGCTTCGATGGGCTTCCTGCACT	1980
Qy	1278	GCGCCATCTCCCTGGTCTCTCTCTGTCATGGACCGGCTGGTGCAGCGATTTCGGCAGTC	1337
Db	1981	GCGCCATCTCCCTGGTCTCTCTCTGTCATGGACCGGCTGGTGCAGCGATTTCGGCAGTC	2040
Qy	1338	GACAGTCATTTGGCCAGTGTGGCAGCTTTCCCTGTGGTGCCGCTGCCACATGCTGT	1397
Db	2041	GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGTGCCGCTGCCACATGCTGT	2100
Qy	1398	CCACAGTGTGGCGGTGGTACAGCTTCAGCCGCTCCACCGGCTTCACTTCTCAGCCC	1457
Db	2101	CCACAGTGTGGCGGTGGTACAGCTTCAGCCGCTCCACCGGCTTCACTTCTCAGCCC	2160
Qy	1458	TGCAGATCTTCCCTACACACTGGCCCTCCCTTACCACCGGGAGAGCAGGTGTTCCCTGC	1517
Db	2161	TGCAGATCTTCCCTACACACTGGCCCTCCCTTACCACCGGGAGAGCAGGTGTTCCCTGC	2220
Qy	1518	CCAAATACGAGGGGACACTGGAGTGTCTAGCAGTACAGCAGCCGTGATGACAGCTTCC	1577
Db	2221	CCAAATACGAGGGGACACTGGAGTGTCTAGCAGTACAGCAGCCGTGATGACAGCTTCC	2280
Qy	1578	TGCAGGCTTCCAGCTGGAGCTCCCTTCCCTTAATGACACAGTGGGTGCTGGAGGAGTG	1637
Db	2281	TGCAGGCTTCCAGCTGGAGCTCCCTTCCCTTAATGACACAGTGGGTGCTGGAGGAGTG	2340
Qy	1638	GCCTGTCTCCACCTCCACCGCTCTGCGGGGCTCTGCTGTGATGTCTCGTACGTG	1697
Db	2341	GCCTGTCTCCACCTCCACCGCTCTGCGGGGCTCTGCTGTGATGTCTCGTACGTG	2400
Qy	1698	TGGTGTGGGTGAGCCACCGAGCCAGGCTGGTTCGGGCGGGGCACTGCTGGACC	1757
Db	2401	TGGTGTGGGTGAGCCACCGAGCCAGGCTGGTTCGGGCGGGGCACTGCTGGACC	2460
Qy	1758	TGCGCATCTGGATAGTCCCTCTGCTGTCCAGGTGGCCCATCCCTGTTTATGGCT	1817
Db	2461	TGCGCATCTGGATAGTCCCTCTGCTGTCCAGGTGGCCCATCCCTGTTTATGGCT	2520
Qy	1818	CCATTGTCCAGCTCAGCAGCTGTCTACTGCTATATGCTGTGCGCGAGGCTGGTC	1877
Db	2521	CCATTGTCCAGCTCAGCAGCTGTCTACTGCTATATGCTGTGCGCGAGGCTGGTC	2580
Qy	1878	TGGTGGCCATTTACTTGTCTACAGAGTATTTGACAGAGGCTGGCCCAATACT	1937
Db	2581	TGGTGGCCATTTACTTGTCTACAGAGTATTTGACAGAGGCTGGCCCAATACT	2640
Qy	1938	CAGCTAGAAACTTCCAGACATTTGGGTGGAGGCTGCTGCTACTGGGTCCAGCTCC	1997
Db	2641	CAGCTAGAAACTTCCAGACATTTGGGTGGAGGCTGCTGCTACTGGGTCCAGCTCC	2700
Qy	1998	CCGCTCTCTGTAGCCCATGGGCTGCCGCTGCCAGTTCCTGTTGCTGCCAAAG	2057
Db	2701	CCGCTCTCTGTAGCCCATGGGCTGCCGCTGCCAGTTCCTGTTGCTGCCAAAG	2760
Qy	2058	TAATGTGGCTCTGTGCTGCCACCTGTGCTGCTGAGTGGCTGCTGCTGAGGCTGGGGC	2117
Db	2761	TAATGTGGCTCTGTGCTGCCACCTGTGCTGCTGAGTGGCTGCTGCTGAGGCTGGGGC	2820
Qy	2118	TGGGCGTCCCTCTCTCTCCAGTCTCTAGGCTGCTGACTGAGGCTTCCAAAG	2177
Db	2821	TGGGCGTCCCTCTCTCTCTCCAGTCTCTAGGCTGCTGACTGAGGCTTCCAAAG	2880
Qy	2178	GGGTTTTCAGTCTGGACTTATACAGGAGGAGGAGGCTTCCATGACACTGGAATCGGG	2237
Db	2881	GGGTTTTCAGTCTGGACTTATACAGGAGGAGGAGGCTTCCATGACACTGGAATCGGG	2940
Qy	2238	GACTCTGAGTGGATTTACCCAGGCTCAGGCTTACAGCTAGCTTCCCTAGTGGAGACAC	2297
Db	2941	GACTCTGAGTGGATTTACCCAGGCTCAGGCTTACAGCTAGCTTCCCTAGTGGAGACAC	3000
Qy	2298	CCTAGAGAGGTTTTCGGAGCTGAATAAAGTCACTCAGTCCCTGGTTCCTCCTCCTAAGC	2357
Db	3001	CCTAGAGAGGTTTTCGGAGCTGAATAAAGTCACTCAGTCCCTGGTTCCTCCTCCTAAGC	3060
Qy	2358	CCCTTAACCTGCAGCTTCTGATGAGTCTTTCGATGGAGTTCCTAGATGAACAC	2417
Db	3061	CCCTTAACCTGCAGCTTCTGATGAGTCTTTCGATGGAGTTCCTAGATGAACAC	3120
Qy	2418	TCCTCATGAGGATTTGAACATATG--ACTTATTTGAGGAGAGTCTCTGAGGGCAAC	2475
Db	3121	TCCTCATGAGGATTTGAACATATG--ACTTATTTGAGGAGAGTCTCTGAGGGCAAC	3180
Qy	2476	ACACAAGAACAGGCTCCCTCAGCCACAGCAGTCTCTTTTGTGATGCCACCCCTCTCT	2535
Db	3181	ACACAAGAACAGGCTCCCTCAGCCACAGCAGTCTCTTTTGTGATGCCACCCCTCTCT	3240
Qy	2536	TACCTTTTATCAGGATGGGCTGTGCTTCTGTTGCTGATGCCATCAGAGACAGGAT	2595
Db	3241	TACCTTTTATCAGGATGGGCTGTGCTTCTGTTGCTGATGCCATCAGAGACAGGAT	3299
Qy	2596	TTAAATATTTAACTTATTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT	2655
Db	3300	TTAAATATTTAACTTATTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT	3359
Qy	2656	GTGTGTCTTAATTTTGGGTAGGCTGGGAGTCCCAACATAGTATTAATTCAGATGAT	2715
Db	3360	GTGTGTCTTAATTTTGGGTAGGCTGGGAGTCCCAACATAGTATTAATTCAGATGAT	3419
Qy	2716	CTGTGTCTTAATTTTGGGTAGGCTGGGAGTCCCAACATAGTATTAATTCAGATGAT	2775
Db	3420	CTGTGTCTTAATTTTGGGTAGGCTGGGAGTCCCAACATAGTATTAATTCAGATGAT	3479
Qy	2776	GCCTAACCCAGGCTTGGAAATCTTACTATCTCCAAATAGTATTAATTCAGATGAT	2835
Db	3480	GCCTAACCCAGGCTTGGAAATCTTACTATCTCCAAATAGTATTAATTCAGATGAT	3539
Qy	2836	CCAAGGTTAGGCTTGTGAAGAGGTAGAGGCTGGGCTTCAAGGCTTCAAGGCTTCCCT	2895
Db	3540	CCAAGGTTAGGCTTGTGAAGAGGTAGAGGCTGGGCTTCAAGGCTTCAAGGCTTCCCT	3599
Qy	2896	AACACCCCTCTCTCTCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	2955
Db	3600	AACACCCCTCTCTCTCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	3659
Qy	2956	TCTAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	3015
Db	3660	TCTAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	3719
Qy	3016	CCCCAACCTTCCCAACAGCTTCCCAACAGCTTCCCAACAGCTTCCCAACAGCTTCC	3075
Db	3720	CCCCAACCTTCCCAACAGCTTCCCAACAGCTTCCCAACAGCTTCCCAACAGCTTCC	3779
Qy	3076	CAAAGTGGGCTTCCCAAGCTTGTGCTATCTAGGCTGGGCTGGGCTGGGCTGGGCTGGG	3135
Db	3780	CAAAGTGGGCTTCCCAAGCTTGTGCTATCTAGGCTGGGCTGGGCTGGGCTGGGCTGGG	3839
Qy	3136	GATCTCACACAGAACTCAGGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3195
Db	3840	GATCTCACACAGAACTCAGGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3899

QY	3196	AGGGGGGTTTAAGTGCCTTCGCTATTATTTACGGGGTGAATAT	3355
Db	3900	AGGGGGGTTTAAGTGCCTTCGCTATTATTTACGGGGTGAATAT	3959
QY	3256	TTTATACGTGAAGTGCAGCAATCAGAGTATTAATGTTATGTCACAAATTAAGGGCTTC	3315
Db	3960	TTTATACGTGAAGTGCAGCAATCAGAGTATTAATGTTATGTCACAAATTAAGGGCTTC	4019

RESULT 17

US-09-780-669-704

; Sequence 704, Application US/097806669

; Patent No. US20020051977A1

; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqun

APPLICANT: Henderson, Robert A

APPLICANT: HEINDELSON, ROBERT D  
APPLICANT: KALOS, MICHAEL D

APPLICANT: Fanger. Garv

: AFFILIANT: Fanger, Gary  
: APPLICANT: Better, Marc: AFFILIANT: KELLER, MARC A  
: APPLICANT: STOLK, JOHN A

: APPLICANT: Day, Craig H.

APPLICANT: Day, Craig H.  
APPLICANT: Vedvick Thomas S

APPLICANT: VEDVICK, Thomas

APPLICANT: Carter, Darrin

APPLICANT: LI, Samuel

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir

APPLICANT: Heppler, William

APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

APPLICANT: Houghton, Raymond L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

; TITLE OF INVENTION: DIAGNOSIS

FILE REFERENCE: 210121.427C24

; CURRENT APPLICATION NUMBER

; CURRENT FILING DATE: 2001-02-09

; NUMBER OF SEQ

; SOFTWARE: Fa

; SEQ ID NO 70

; LENGTH: 4034

; TYPE: DNA

Query Match 75.8%; Score 2585.4; DB 10; Length 4034;  
Best Local Similarity 82.4%; Pred. No. 0;  
Matches 3325; Conservative 0; Mismatches 1; Indels 709; Gaps 3;

4	AACAGCGCTGACCGCTGCGTCCGGGTGACAGCCGCGGCTCGGCCAGGATCTGATG	63
Qy		
1	AACAGCGCTGACCGCTGCGTCCGGGTGACAGCCGCGGCTCGGCCAGGATCTGATG	60
Db		
64	ATGAGAGCTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG	123
Qy		
61	ATGAGAGCTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG	120
Db		
124	CTGAGCCGGCACCAAGGGCTGGCAGAAATGGCGCTCGCTGATTCTAGCGAGTTGGC	183
Qy		
121	CTGAGCCGGCACCAAGGGCTGGCAGAAATGGCGCTGGCTGATTCTAGCGAGTTGGC	180
Db		
184	GGCAGCAGGAGGAGGCCGCGACGCTTCGTGGAGCAGCCGAGACGAGATTCTGGAG	243
Qy		
181	GGCAGCAGGAGGAGGCCGCGACGCTTCGTGGAGCAGCCGAGACGAGATTCTGGAG	240
Db		
244	TGCTTGAAACGGCCCTTGAGCCCTACCCCGCTGGCCCACTATGTGTCCAGAGCGCTGTGGGT	303
Qy		

QY	1240	-----	1239
Db	1381	CGGCTTCTCATGGGTGGAACATCTCTGTTCCGGTTTCAGGAAGCCCTCTGGGTGCTC	1440
QY	1240	-----	1239
Db	1441	TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGAAAGCGAGCTTAATCAA	1500
QY	1240	-----	1239
Db	1501	AGTCTAGAGGAGTGGAGGAGTTAAGCTGGATTTCAGATCTGCTGGTTCCAGCGCAG	1560
QY	1240	-----	1239
Db	1561	TGTGCCCTCTGCTCCCCAACAGACTTTCNAATAATCTCACAGCCCTTTCAGCTCAGG	1620
QY	1240	-----	1239
Db	1621	CGTCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGCTTTGTGTTCCCTCTCACCCGCT	1680
QY	1240	-----	1239
Db	1681	GTCTCTCACAGTCTGAGACTCCAGGAACCTTTCAGACTACCTTCTCTGCTTTCAGCAAG	1740
QY	1240	-----	1239
Db	1741	GGCGTGTCCCAATCTCTCTGAGGCTCAGTGGGAAGAACCTAGACTCCCATCTGTAGAGTA	1800
QY	1240	-----	1239
Db	1801	GAAGGGAAGGTGCTGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC	1860
QY	1240	-----	1239
Db	1861	CTGTGTTCCGCTTCTCTATCTCTCTGAGACTGCTCCGACCTTCCCTCCAGGCTCTGT	1920
QY	1240	-----	1277
Db	1921	CTGATGCCCTCTCCCTCTGACAGGCTTCGGATGGGACGCTGGGGCTGTTCTCTGCACT	1980
QY	1278	CGCCATCTCCCTGGTCTCTCTCTGATGACCGGCTGTGTCAGCGATTGGGCACCTC	1337
Db	1981	CGCCATCTCCCTGGTCTCTCTCTGATGACCGGCTGTGTCAGCGATTGGGCACCTC	2040
QY	1338	GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGTGCCACATGCTGT	1397
Db	2041	GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGTGCCACATGCTGT	2100
QY	1398	CCACAGTGTGGCGGTGTCACAGCTTTCAGCGCCCTCACCGGGTTTCACTTCTCAGCCG	1457
Db	2101	CCACAGTGTGGCGGTGTCACAGCTTTCAGCGCCCTCACCGGGTTTCACTTCTCAGCCG	2160
QY	1458	TGCAGATCTTCCCTACACACTGGCTCCCTCTACACCGGGAGAGAGGTTTCTCTG	1517
Db	2161	TGCAGATCTTCCCTACACACTGGCTCCCTCTACACCGGGAGAGAGGTTTCTCTG	2220
QY	1518	CCAAATACGAGGGGACACTGGAGGTCTAGCAGTGGAGACGCTGATGACAGCTTCC	1577
Db	2221	CCAAATACGAGGGGACACTGGAGGTCTAGCAGTGGAGACGCTGATGACAGCTTCC	2280
QY	1578	TGCCAGGCTTAAGCTTGGAGCTTCCCTTCCCTTAATGACACGCTGGGTGCTGGAGGAGTG	1637
Db	2281	TGCCAGGCTTAAGCTTGGAGCTTCCCTTCCCTTAATGACACGCTGGGTGCTGGAGGAGTG	2340
QY	1638	GCCTGTCTCCACCTTCCACCGCGCTCTGCGGGGCTCTGCTGTGATGCTCTCGTACGTG	1697
Db	2341	GCCTGTCTCCACCTTCCACCGCGCTCTGCGGGGCTCTGCTGTGATGCTCTCGTACGTG	2400
QY	1698	TGTTGGTGGGTGAGCCACCGAGGCTTGGGTGCTGGGGCCGGGGCATCTGCTGGACC	1757
Db	2401	TGTTGGTGGGTGAGCCACCGAGGCTTGGGTGCTGGGGCCGGGGCATCTGCTGGACC	2460
QY	1758	TCGCCATCTGATAGTGCCTTCTCTGCTGCCAGGTGGCCCATCTCTGTTATGGGCT	1817
Db	2461	TCGCCATCTGATAGTGCCTTCTCTGCTGCCAGGTGGCCCATCTCTGTTATGGGCT	2520
QY	1818	CAATGTCCAGTCTAGCCAGTCTGTCACTGCTATATGGTGTCTGCGCAGGCTGGGTC	1877
Db	2521	CAATGTCCAGTCTAGCCAGTCTGTCACTGCTATATGGTGTCTGCGCAGGCTGGGTC	2580
QY	1878	TGTCGCCATTTACTTTGCTACAGGTAGTATTTGACAAGAGGAGCTTGGCCAAATACT	1937
Db	2581	TGTCGCCATTTACTTTGCTACAGGTAGTATTTGACAAGAGGAGCTTGGCCAAATACT	2640
QY	1938	CAGCGTAGAAAATTTCCAGCACATTTGGGGTGGAGGCTTGCCTCACTGGGTCCAGCTCC	1997
Db	2641	CAGCGTAGAAAATTTCCAGCACATTTGGGGTGGAGGCTTGCCTCACTGGGTCCAGCTCC	2700
QY	1998	CGCTCTCTGTAGCCCCATGGGGTGGCGGCTGGCCGCCAGTTTCTGTTGCTGCGCAAG	2057
Db	2701	CGCTCTCTGTAGCCCCATGGGGTGGCGGCTGGCCGCCAGTTTCTGTTGCTGCGCAAG	2760
QY	2058	TAATCTGCTCTCTGCTGCCACCTGCTGCTGAGGTGCGTGTGCTGACAGTGGGGG	2117
Db	2761	TAATCTGCTCTCTGCTGCCACCTGCTGCTGAGGTGCGTGTGCTGACAGTGGGGG	2820
QY	2118	TGGGGCTCCCTCTCTCTCCCGAGTCTCTAGGGCTGCTGTGCTGAGGCTTCCAG	2177
Db	2821	TGGGGCTCCCTCTCTCTCTCCCGAGTCTCTAGGGCTGCTGTGCTGAGGCTTCCAG	2880
QY	2178	GGGGTTTCACTGTGACTTATACAGGAGGCGGAGGCTTCCATGCTGCTGCTGCTGCTG	2237
Db	2881	GGGGTTTCACTGTGACTTATACAGGAGGCGGAGGCTTCCATGCTGCTGCTGCTGCTG	2940
QY	2238	GACTCTGAGGTGGATTACCCAGGCTCAGGGTTTACAGCTAGCTTCTAGTTGACACACA	2297
Db	2941	GACTCTGAGGTGGATTACCCAGGCTCAGGGTTTACAGCTAGCTTCTAGTTGACACACA	3000
QY	2298	CTAGAGAGGGTTTTTGGGAGCTGAATAAATCACTAGCTACCTGCTGCTTCCATCTTAAGC	2357
Db	3001	CTAGAGAGGGTTTTTGGGAGCTGAATAAATCACTAGCTACCTGCTGCTTCCATCTTAAGC	3060
QY	2358	CCCTTAACCTGCAGCTTCTGTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAACAC	2417
Db	3061	CCCTTAACCTGCAGCTTCTGTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAACAC	3120
QY	2418	TCCTCCATGGGATTTGAACATATGAAAGTTATTTAGGGGAGAGTCTCTGAGGGGCAAC	2475
Db	3121	TCCTCCATGGGATTTGAACATATGAAAGTTATTTAGGGGAGAGTCTCTGAGGGGCAAC	3180
QY	2476	ACACAAGAACACAGGTCCCTCAGCCACAGCACTCTCTTTTGTGCTGATCCACCCCTCT	2535
Db	3181	ACACAAGAACACAGGTCCCTCAGCCACAGCACTCTCTTTTGTGCTGATCCACCCCTCT	3240
QY	2536	TACCTTTTATCAGGATGTGGCTGTGTTGCTTCTGTTGCCATCAGAGACACAGGAT	2595
Db	3241	TACCTTTTATCAGGATGT - GCCTGTTGGTCTCTCTGTTGCCATCAGAGACACAGGAT	3299
QY	2596	TTAATATTTAATTTATTTATTTACAAAGTAGAAGGAATCCATTTGCTGCTTTCTGT	2655
Db	3300	TTAATATTTAATTTATTTATTTAACAAGTAGAAGGAATCCATTTGCTGCTTTCTGT	3359
QY	2656	GTGTGCTCTAATTTTGGGTAGGTGGGGATCCCAACATCAGGTCCCTCTGATAGTAG	2715
Db	3360	GTGTGCTCTAATTTTGGGTAGGTGGGGATCCCAACATCAGGTCCCTCTGATAGTAG	3419
QY	2716	CTGGTCAATTTGGCTGATCATTTGCCAGATCTTCTCTCTCTGGGGTCTGGCCCCCAAA	2775
Db	3420	CTGGTCAATTTGGCTGATCATTTGCCAGATCTTCTCTCTCTGGGGTCTGGCCCCCAAA	3479
QY	2776	GCCTAACCCAGGACCTTGGAAATTTCTACTATCCCAATGATAATTCAAATGCTGTAC	2835
Db	3480	GCCTAACCCAGGACCTTGGAAATTTCTACTATCCCAATGATAATTCAAATGCTGTAC	3539
QY	2836	CCAAGTTAGGCTTGTGAAGGAAGGTAGAGGTGGGGCTTTCAGGTCTTCAACGGCTTCCCT	2895

```
Db 3540 CCAGGTTAGGTTGAGGAGGTTAGAGGTTGGGCTTCAGGTTCTCAACGGCTTCCCT 3599
Qy 2896 AACACCCCTCTCTCTTTGGCCAGCCTGGTTCCGCCCACTTCCACTCCCTCTACTCTC 2955
Db 3600 AACACCCCTCTCTCTTTGGCCAGCCTGGTTCCGCCCACTTCCACTCCCTCTACTCTC 3659
Qy 2956 TCTAGGACTGGGTGATGAGGACTGCGCCAAATTTCCCTTACCCCACTTCCCTCTA 3015
Db 3660 TCTAGGACTGGGTGATGAGGACTGCGCCAAATTTCCCTTACCCCACTTCCCTCTA 3719
Qy 3016 CCCCCAACTTCCCAACAGCTCCACACCCCTGTTGGAGCTACTGAGGAGCCAGAGCA 3075
Db 3720 CCCCCAACTTCCCAACAGCTCCACACCCCTGTTGGAGCTACTGAGGAGCCAGAGCA 3779
Qy 3076 CAAAGTGGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAAGATATATCTGTGCTGG 3135
Db 3780 CAAAGTGGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAAGATATATCTGTGCTGG 3839
Qy 3136 GAATCTCACAGAACTCAGGAGCAGCCCTGCTGAGCTAAGGAGGCTTATCTCTC 3195
Db 3840 GAATCTCACAGAACTCAGGAGCAGCCCTGCTGAGCTAAGGAGGCTTATCTCTC 3899
Qy 3196 AGGGGGGTTTAAAGTCCCGTTTGCATATATGCTCTTATTTAGCGGGTGAATAT 3255
Db 3900 AGGGGGGTTTAAAGTCCCGTTTGCATATATGCTCTTATTTAGCGGGTGAATAT 3959
Qy 3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAGGCTTC 3315
Db 3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAGGCTTC 4019
Qy 3316 TTTATGTTTAAAAA 3330
Db 4020 TTTATGTTTAAAAA 4034

RESULT 18
US-09-822-827-704
; Sequence 704, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 704
; LENGTH: 4034
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-704

Query Match 75.8%; Score 2585.4; DB 10; Length 4034;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 3325; Conservative 0; Mismatches 1; Indels 709; Gaps 3;

Qy 4 AACAGCCTGACGCGCTGGTTCGGGTGACAGCGCGCGCTCGGCCAGGATCTGAGTG 63
Db 1 AACAGCCTGACGCGCTGGTTCGGGTGACAGCGCGCGCTCGGCCAGGATCTGAGTG 60
Qy 64 ATGAGACGTGTCCTCCACTGAGTGTCCTCCACAGCAGCAGGTTGAGCATGGCTGAGAAG 123
Db 61 ATGAGACGTGTCCTCCACTGAGTGTCCTCCACAGCAGCAGGTTGAGCATGGCTGAGAAG 120
Qy 124 CTGGACCGGCACCAAGGCTGGCAGAAATGGCGCCCTGGCTGATTCCTAGGAGTTGGC 183
Db 121 CTGGACCGGCACCAAGGCTGGCAGAAATGGCGCCCTGGCTGATTCCTAGGAGTTGGC 180
Qy 184 GGCAGCAGGAGGAGGAGCGCGCAGCTTCTGGAGCAGAGCGCGCAGGAGCTTCTGGAG 243
Db 185 GGCAGCAGGAGGAGGAGCGCGCAGCTTCTGGAGCAGAGCGCGCAGGAGCTTCTGGAG 243
```

```
Db 181 GGCAGCAGGAGGAGGAGCGCGCAGCTTCTGGAGCAGAGCGCGCAGGAGCTTCTGGAG 240
Qy 244 TGCCCTGAACGGCCCGCTTACCCCTTACCCGCTGCGCCCACTATGCTCCAGAGGCTGTGGGT 303
Db 241 TGCCCTGAACGGCCCGCTTACCCCTTACCCGCTGCGCCCACTATGCTCCAGAGGCTGTGGGT 300
Qy 304 GAGCGGCTGCTGGGACCGGAAAGCCAGCTTCTGCTGGTCAACCTGCTAACTTGG 363
Db 301 GAGCGGCTGCTGGGACCGGAAAGCCAGCTTCTGCTGGTCAACCTGCTAACTTGG 360
Qy 364 CTTGAGGTTGTTTGGCGCAGCATCACTATGTCGCCCTCTGCTGCTGGAAGTGG 423
Db 361 CTTGAGGTTGTTTGGCGCAGCATCACTATGTCGCCCTCTGCTGCTGGAAGTGG 420
Qy 424 GGTAGAGGAGAAATTCATGACCATGCTGGGCACTTGGTCCAGTGTGGGCTGTGCTG 483
Db 421 GGTAGAGGAGAAATTCATGACCATGCTGGGCACTTGGTCCAGTGTGGGCTGTGCTG 480
Qy 484 TGTCCGCTCCTAGGCTCAGCAGTGAACACTGGCTGGAGCTATGCGCGCGCGCGGCC 543
Db 481 TGTCCGCTCCTAGGCTCAGCAGTGAACACTGGCTGGAGCTATGCGCGCGCGCGGCC 540
Qy 544 CTTGCTGAGGAGGCTGCTGTCGCCGATCCAGGCGCTTGGAGCTGGAGCTGGCTCATCT 603
Db 541 CTTGCTGAGGAGGCTGCTGTCGCCGATCCCTGGGCACTTCTGCTGAGGCTCTTCTCAT 600
Qy 604 CTGCTAGCAGGAGGCTGCTGTCGCCGATCCAGGCGCTTGGAGCTGGAGCTGGCTCATCT 663
Db 601 CTGCTAGCAGGAGGCTGCTGTCGCCGATCCAGGCGCTTGGAGCTGGAGCTGGCTCATCT 660
Qy 664 GGGGCTGGGCTGCTGAGCTTCTGTCGCCAGTGTGCTTCACTCCACTGGAGGCGCTGCT 723
Db 661 GGGGCTGGGCTGCTGAGCTTCTGTCGCCAGTGTGCTTCACTCCACTGGAGGCGCTGCT 720
Qy 724 CTCTGAGCTTCTCGGAGCAGGACACTGCTGCGCAGGCTTACTGCTGCTATGCTTCAT 783
Db 721 CTCTGAGCTTCTCGGAGCAGGACACTGCTGCGCAGGCTTACTGCTGCTATGCTTCAT 780
Qy 784 GATCAGCTTTGGGGCTGCTGCGGCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843
Db 781 GATCAGCTTTGGGGCTGCTGCGGCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 844 CTTGGGCGCTTACTGCGGACCGAGGAGTGGCTTCTTTGGGCTGCTGCTGCTGCTGCTT 903
Db 841 CTTGGGCGCTTACTGCGGACCGAGGAGTGGCTTCTTTGGGCTGCTGCTGCTGCTGCTT 900
Qy 904 CTTGAGCTGCTGAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963
Db 901 CTTGAGCTGCTGAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy 964 GGCAGCAGGAGGCTGTCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023
Db 961 GGCAGCAGGAGGCTGTCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1024 CTTGGCTTTCCGNACTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1083
Db 1021 CTTGGCTTTCCGNACTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1084 GCGCGCAGCCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143
Db 1081 GCGCGCAGCCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy 1144 CTTGAGCTGCTTTTACACGATTTCTGCTGGGAGGAGGCTGTACCAGGCTGCTGCTGCTGCT 1203
Db 1141 CTTGAGCTGCTTTTACACGATTTCTGCTGGGAGGAGGCTGTACCAGGCTGCTGCTGCTGCT 1200
Qy 1204 TGAGCGGCGCAGGCGCGGAGACACTATGATGA----- 1239
Db 1201 TGAGCGGCGCAGGCGCGGAGACACTATGATGAAGGCTTGGAGCTTGGAGCAGCAG 1260
Qy 1240 ----- 1239
Db 1261 AGGCTGCTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
```

Qy	1240	-----	1239
Db	1321	TCCTCATCTGGCCCGAGCTCTCTGTGAGAAAGTGGGATGGACCCCATCTGCATACA	1380
Qy	1240	-----	1239
Db	1381	CGSCTTCTCATGGGTGTGAACATCTCTGTTCCGGTTTCAGGAAGCCCTCTGGCTGCTC	1440
Qy	1240	-----	1239
Db	1441	TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGCACAGAAGGAAGCGGAGCTTATTCAA	1500
Qy	1240	-----	1239
Db	1501	AGTCTAGAGGAGTGAGGAGTTAAGGCTTGAATTCAGATCTGCTTGGTTCCAGCCGAG	1560
Qy	1240	-----	1239
Db	1561	TGTGCCCTCTGTCCCCCAACGACTTTCAAATAATCTCACAGCCGCTTCCAGCTCAGG	1620
Qy	1240	-----	1239
Db	1621	CGTCTAGAAGCGTCTTGAAGCCTATGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCT	1680
Qy	1240	-----	1239
Db	1681	GTCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCTCTGCTTTCAGCAAGG	1740
Qy	1240	-----	1239
Db	1741	GGCGTTGCCACATTCCTCTGAGGTCAGTGGAGAACCCTAGACTCCCATTTGCTAGAGTA	1800
Qy	1240	-----	1239
Db	1801	GAAAGGGAAGGTGCTGGGGAGCAGGGCTGTCCACAGCAGCTCTCGTGCAGCAGGTAC	1860
Qy	1240	-----	1239
Db	1861	CTGTGTTCCGCCCTTCTATCTCCCTGAGACTGCTCCGACCCCTTCCCTCCCAGGCTCTGT	1920
Qy	1240	-----AGGCGTTCCGATGGCAGCCTGGGGCTGTTCCTGCAGT	1277
Db	1921	CTGATGGCCCTCTCCCTCTGCAGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT	1980
Qy	1278	GGCCATCTCCCTGCTCTCTCTGTCATGACCGGCTGTGTGACGATTCGGCACTC	1337
Db	1981	GCCCATCTCCCTGGTCTCTCTGTCATGACCGGCTGTGTGACGATTCGGCACTC	2040
Qy	1338	GAGCAGTCTATTGGCCAGTGGCAGCTTCCCTGTGCTGCCGTGCCAGTCATGCTCT	1397
Db	2041	GAGCAGTCTATTGGCCAGTGGCAGCTTCCCTGTGCTGCCGTGCCAGTCATGCTCT	2100
Qy	1398	CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTTCACTTCTCAGCCC	1457
Db	2101	CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTTCACTTCTCAGCCC	2160
Qy	1458	TGCAGATCTGCCCTACACACTGGCCCTCCCTCTACACCGGGAGAGCAGGTGTTCCCTGC	1517
Db	2161	TGCAGATCTGCCCTACACACTGGCCCTCCCTCTACACCGGGAGAGCAGGTGTTCCCTGC	2220
Qy	1518	CCAAATACCGAGGGGACACTGGAGTGTCTAGCAGTCAGGACAGCCCTGATGACCACTTCC	1577
Db	2221	CCAAATACCGAGGGGACACTGGAGTGTCTAGCAGTCAGGACAGCCCTGATGACCACTTCC	2280
Qy	1578	TGCCAGGCCCTAAGCCTTGGAGTCTCCCTTCCCTAATGGACACGTTGGTGTGGAGGCAGTG	1637
Db	2281	TGCCAGGCCCTAAGCCTTGGAGTCTCCCTTCCCTAATGGACACGTTGGTGTGGAGGCAGTG	2340
Qy	1638	GCCTGTCCCACTCCACCGGCGCTCTGGGGGCGCTCTGCCTGTGTATGTCCTCGTACGTG	1697
Db	2341	GCCTGTCCCACTCCACCGGCGCTCTGGGGGCGCTCTGCCTGTGTATGTCCTCGTACGTG	2400

QY	1698	TGTTGGTGGGTGAGCCACCAGAGGCCAGGTGTTCGGGCGGGGCATCTGCCGTGGACC	1755
DB	2401	TGTTGGTGGGTGAGCCACCAGGCCAGGTGTTCGGGCGGGGCATCTGCCGTGGACC	2460
QY	1758	TCGCCATCCTGGATAGTGGCTTCCTGCTGTCCAGGTGGCCCCATCCCTGTTTATGGGCT	1817
DB	2461	TCGCCATCCTGGATAGTGGCTTCCTGCTGTCCAGGTGGCCCCATCCCTGTTTATGGGCT	2520
QY	1818	CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGTTGTCTGCCGCAAGGCTCTGGTC	1877
DB	2521	CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGTTGTCTGCCGCAAGGCTCTGGTC	2580
QY	1878	TGTTGCGCAATTTACTTTGTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT	1937
DB	2581	TGTTGCGCAATTTACTTTGTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT	2640
QY	1938	CAGCGTAGAAAACTTCCAGCACATTTGGGTGGAGGGCTTCGCTCACTGGTCCCACTCC	1997
DB	2641	CAGCGTAGAAAACTTCCAGCACATTTGGGTGGAGGGCTTCGCTCACTGGTCCCACTCC	2700
QY	1998	CCGCTCCTGTTTAGCCCCATGGGCTGCCGGCTGGCGCCAGTTTCTGTTGCTGCCAAAG	2057
DB	2701	CCGCTCCTGTTTAGCCCCATGGGCTGCCGGCTGGCGCCAGTTTCTGTTGCTGCCAAAG	2760
QY	2058	TAATGTGGCTCTCTGTGCCACCTGTGTGCTGAGGTGCTAGCTGACAGCTGGGGGC	2117
DB	2761	TAATGTGGCTCTCTGTGCCACCTGTGTGCTGAGGTGCTAGCTGACAGCTGGGGGC	2820
QY	2118	TGGGGCTCCCTCTCTCTCTCCAGTCTCTTAGGGCTGCCTGACTGGAGGCTTCCAAG	2177
DB	2821	TGGGGCTCCCTCTCTCTCTCCAGTCTCTTAGGGCTGCCTGACTGGAGGCTTCCAAG	2880
QY	2178	GGGTTTTCAGTCTGAGCTTATACAGGAGGCCAGAGGGCTCCATGCACTTGAATCGGG	2237
DB	2881	GGGTTTTCAGTCTGAGCTTATACAGGAGGCCAGAGGGCTCCATGCACTTGAATCGGG	2940
QY	2238	GACTCTCGAGGTGGATTACCCAGCTCAGGGTTAAACAGCTAGCTTCTAGTTGAGACACA	2297
DB	2941	GACTCTCGAGGTGGATTACCCAGCTCAGGGTTAAACAGCTAGCTTCTAGTTGAGACACA	3000
QY	2298	CCTAGACAAAGGTTTTTGGAGCTGAATAAAGTCAAGTCACTGCTGGTTTCCCATCTCTAAGC	2357
DB	3001	CCTAGACAAAGGTTTTTGGAGCTGAATAAAGTCAAGTCACTGCTGGTTTCCCATCTCTAAGC	3060
QY	2358	CCCTTAACTGTCAGCTTCGTTTAAATGTAGCTCTTGATGGAGGTTTCTAGGATGAACAC	2417
DB	3061	CCCTTAACTGTCAGCTTCGTTTAAATGTAGCTCTTGATGGAGGTTTCTAGGATGAACAC	3120
QY	2418	TCCTCCATGGATTTGAACATATG--ACTTATTTGTAGGGGAAGTCTCTGAGGGGCAAC	2475
DB	3121	TCCTCCATGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGTCTCTGAGGGGCAAC	3180
QY	2476	ACACAAAGCCAGTCCCTCAGGCCACAGCACTGCTTTTTCGTGATCCACCCCTCT	2535
DB	3181	ACACAAAGCCAGTCCCTCAGGCCACAGCACTGCTTTTTCGTGATCCACCCCTCT	3240
QY	2536	TACCTTTTATCAGGATCTGGCCTGTTGGTCTCTCTGTGTGCATCACAGAGACACAGGCAT	2595
DB	3241	TACCTTTTATCAGGATCT--GCCGTGGTCTCTCTGTGTGCATCACAGAGACACAGGCAT	3299
QY	2596	TTAAATATTTAACTTATTTATTTAAACAACTAGAGGGAATCCATTTGCTAGCTTTTCTGT	2655
DB	3300	TTAAATATTTAACTTATTTATTTAAACAACTAGAGGGAATCCATTTGCTAGCTTTTCTGT	3359
QY	2656	GTTCGTCTCTAAATATTTGGGTAGGTGGGGATCCCCAACAACTCAGTCCCTCGAGATAG	2715
DB	3360	GTTCGTCTCTCTAAATATTTGGGTAGGTGGGGATCCCCAACAACTCAGTCCCTCGAGATAG	3419
QY	2716	CTGTCATTTGGGCTGATCATTTGCCAAGTCTTCTCTCTGGGCTCTGGCCCCCAAAAT	2775
DB	3420	CTGTCATTTGGGCTGATCATTTGCCAAGTCTTCTCTCTGGGCTCTGGCCCCCAAAAT	3479
QY	2776	GCCTTAACCCAGGACCTTTGGAAATTTCTACTCATCCCAAATGATAATTTCCAAATGCTCTTAC	2835



Db 3480 GCCTAACCCAGGACCTTGGAAATTCCTACTATCCCAATGATATTCAAATGCTGTAC 3539  
Qy 2836 CCAAGGTTAGGCTTTGAAGGAAGGTAGAGGGTGGGCTTTCAGGTCTCTCAACGGCTTCCCT 2895  
Db 3540 CCAAGGTTAGGCTTTGAAGGAAGGTAGAGGGTGGGCTTTCAGGTCTCTCAACGGCTTCCCT 3599  
Qy 2896 AACCAACCCCTTCTCTTTGGCCAGCCTGCTTCCACCTTCCACCTTCCACCTTCTATCTC 2955  
Db 3600 AACCAACCCCTTCTCTTTGGCCAGCCTGCTTCCACCTTCCACCTTCCACCTTCTATCTC 3659  
Qy 2956 TCTAGGACTGGCTGATGAGGCACTGCCCAAAATTTCCCTACCCCTTCCCTTCCCTA 3015  
Db 3660 TCTAGGACTGGCTGATGAGGCACTGCCCAAAATTTCCCTACCCCTTCCCTTCCCTA 3719  
Qy 3016 CCCCACAACTTTCCCCACAGCTCCACAAACCTGTTTGGAGCTACTGAGGACCAAGCA 3075  
Db 3720 CCCCACAACTTTCCCCACAGCTCCACAAACCTGTTTGGAGCTACTGAGGACCAAGCA 3779  
Qy 3076 CAAAGTGGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTCTGG 3135  
Db 3780 CAAAGTGGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTCTGG 3839  
Qy 3136 GAATCTCACAGAACTCAGGAGCACCCCTGCTGAGCTAAGGAGGCTTTATCTCTC 3195  
Db 3840 GAATCTCACAGAACTCAGGAGCACCCCTGCTGAGCTAAGGAGGCTTTATCTCTC 3899  
Qy 3196 AGGGGGGTTTAAAGTGGCTTTGCAATTAATGCTCTTATTTAGCGGGTGAATAT 3255  
Db 3900 AGGGGGGTTTAAAGTGGCTTTGCAATTAATGCTCTTATTTAGCGGGTGAATAT 3959  
Qy 3256 TTTTACTGTAAGTGAGCAATCAGAGTATATGTTTATGTTGACAAATTAAGGCTTTC 3315  
Db 3960 TTTTACTGTAAGTGAGCAATCAGAGTATATGTTTATGTTGACAAATTAAGGCTTTC 4019  
Qy 3316 TTATATGTTTAAAA 3330  
Db 4020 TTATATGTTTAAAA 4034

## RESULT 19

US-10-012-896-703  
; Sequence 703, Application US/10012896  
; Publication No. US2002018251A1

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Mesgher, Madeleine Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C27  
; CURRENT APPLICATION NUMBER: US/10/012,896

; CURRENT FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 1011  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 703  
; LENGTH: 2904  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-012-896-703

Query Match 64.4%; Score 2196.4; DB 9; Length 2904;

Best Local Similarity 88.2%; Pred. No. 0;

Matches 2560; Conservative 0; Mismatches 1; Indels 343; Gaps 2;

Qy 770 GTCATATGCTTCATGATCAGTCTTGGGGCTCCCTGGGGCTACCTCTGCTGCCATTGAC 829  
Db 1 GTCATATGCTTCATGATCAGTCTTGGGGCTCCCTGGGGCTACCTCTGCTGCCATTGAC 60  
Qy 830 TGGGACACAGTCCCTGGCCCCCTACCTGGGACCCAGGAGGAGTGCCTCTTTGGCCTG 889  
Db 61 TGGGACACAGTCCCTGGCCCCCTACCTGGGACCCAGGAGGAGTGCCTCTTTGGCCTG 120  
Qy 890 CTCACCTCATCTTCCTACACTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGCAGCG 949  
Db 121 CTCACCTCATCTTCCTACACTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGCAGCG 180  
Qy 950 CTGGGGCCACAGGAGCAGAGGGCTGTCGGGCCCTCTCTTGTGCCCCACACTGCTGT 1009  
Db 181 CTGGGGCCACAGGAGCAGAGGGCTGTCGGGCCCTCTCTTGTGCCCCACACTGCTGT 240  
Qy 1010 CCATGCGGGCCCGCTTGGCTTTCGGGAACCTGCGGCCCTCTTCCCGGCTGCACCG 1069  
Db 241 CCATGCGGGCCCGCTTGGCTTTCGGGAACCTGCGGCCCTCTTCCCGGCTGCACCG 300  
Qy 1070 CTGTGCTGCCGATGCGCCGACCCCTGCGGCCGCTTTCGTGGCTGAGTGTGACAGTGG 1129  
Db 301 CTGTGCTGCCGATGCGCCGACCCCTGCGGCCGCTTTCGTGGCTGAGTGTGACAGTGG 360  
Qy 1130 ATGGCACTCATGACCTTACGCTGTTTTACAGGATTTGCTGGGGAGGGCTGTACAGG 1189  
Db 361 ATGGCACTCATGACCTTACGCTGTTTTACAGGATTTGCTGGGGAGGGCTGTACAGG 420  
Qy 1190 GCGTGCCGAGAGCTGAGCGGGGACCGAGGCCGAGACACTATGATGAA----- 1240  
Db 421 GCGTGCCGAGAGCTGAGCGGGGACCGAGGCCGAGACACTATGATGAAAGGGCT 480  
Qy 1241 ----- 1240  
Db 481 CTGGCTGCTCTAGAGTCTGTATCAGAGTCTGTTGCCCTTTCACAGAAGGAAGCGGA 540  
Qy 1241 ----- 1240  
Db 541 GCTTATTCAAAGTCTAGAGGAGTGGAGGTTAAGGCTGGATTTTCAGATCTGCTGGTT 600  
Qy 1241 ----- 1240  
Db 601 CCAGCGCAGTGTGCCCTCTGCTCCCCAACGACTTTCACAAATATCTCACCAGCGCCTT 660  
Qy 1241 ----- 1240  
Db 661 CCAGCTCAGGCGTCTTAGAAGGCTTTGAAGGCTATGGCAGCTGTCTTGTGTTCCCTC 720  
Qy 1241 ----- 1240  
Db 721 TCACCGGCTGTCTTCACAGCTGAGACTCCCAGGAACCTTCAGACTACCTTCTCTGCC 780  
Qy 1241 -----GGGTTTCGATGGGAGCTTGGGCTGT 1268  
Db 781 TTCAGCAAGGGGTTGCCACATCTCTGAGGGGCTTCGGATGGCAGCTTGGGCTGT 840  
Qy 1269 TCTGCAGTGCAGCATCTCCCTGGTCTCTCTCTGTATGACCGGCTGTGTGACGAGCAT 1328  
Db 841 TCTGCAGTGCAGCATCTCCCTGGTCTCTCTCTGTATGACCGGCTGTGTGACGAGCAT 900



QY 1329 TCGGCACCTGACAGCTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGTCCGCTGCCA 1388  
DB 901 TCGGCACCTGACAGCTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGTCCGCTGCCA 960  
QY 1389 CATGCCCTGCCACAGTGTGGCGGTGGTGCAGCTTACGCCCTCACCAGGTTCACT 1448  
DB 961 CATGCCCTGCCACAGTGTGGCGGTGGTGCAGCTTACGCCCTCACCAGGTTCACT 1020  
QY 1449 TCTCAGCCCTGACAGTCTCCCTACACACTGCTCCCTCTACACCGGGAGAGCAGG 1508  
DB 1021 TCTCAGCCCTGACAGTCTCCCTACACACTGCTCCCTCTACACCGGGAGAGCAGG 1080  
QY 1509 TGTCTCTGCCAAATACCGAGGGGACACTGAGGTGCTAGCAGTGGAGACAGCTGATGA 1568  
DB 1081 TGTCTCTGCCAAATACCGAGGGGACACTGAGGTGCTAGCAGTGGAGACAGCTGATGA 1140  
QY 1569 CCAGCTTCTGCCAGGCCCTAAGCCTGAGCTCCCTTCCCTAATGGACAGTGGGTGCTG 1628  
DB 1141 CCAGCTTCTGCCAGGCCCTAAGCCTGAGCTCCCTTCCCTAATGGACAGTGGGTGCTG 1200  
QY 1629 GAGGCAGTGGCTGCTCCACCTCCACCCGCGCTCTGCGGGGCTCTGCCTGTGATGCT 1688  
DB 1201 GAGGCAGTGGCTGCTCCACCTCCACCCGCGCTCTGCGGGGCTCTGCCTGTGATGCT 1260  
QY 1689 CCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1748  
DB 1261 CCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1320  
QY 1749 GCCTGACCTGCGCATCTCGATGATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1808  
DB 1321 GCCTGACCTGCGCATCTCGATGATGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
QY 1809 TTATGGGCTCCATTTGCCAGCTCAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1868  
DB 1381 TTATGGGCTCCATTTGCCAGCTCAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
QY 1869 GCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1928  
DB 1441 GCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1929 CCNAATACCTCAGGTAGAAAATTCAGCACATTTGGGTGGAGGCTGCTGCTGCTGCTGCTG 1988  
DB 1501 CCNAATACCTCAGGTAGAAAATTCAGCACATTTGGGTGGAGGCTGCTGCTGCTGCTGCTG 1560  
QY 1989 CCAGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2048  
DB 1561 CCAGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
QY 2049 CTGCCAAAGTAATGTGGCTCTGCTGCTGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 2108  
DB 1621 CTGCCAAAGTAATGTGGCTCTGCTGCTGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
QY 2109 GCTGGGGGCTGGGGGCTCCCT 2168  
DB 1681 GCTGGGGGCTGGGGGCTCCCT 1740  
QY 2169 CTTCTCAAGGGGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2228  
DB 1741 CTTCTCAAGGGGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
QY 2229 GAATGGGGGACTCTGCAGGTGGATTTACCGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288  
DB 1801 GAATGGGGGACTCTGCAGGTGGATTTACCGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860  
QY 2289 TGAGACACACTAGAGAGGGTTTTGGAGCTGAATAAATCAGTCACTGCTGTTTCCCA 2348  
DB 1861 TGAGACACACTAGAGAGGGTTTTGGAGCTGAATAAATCAGTCACTGCTGTTTCCCA 1920  
QY 2349 TCTCTAAGCCCTTACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2408  
DB 1921 TCTCTAAGCCCTTACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980  
QY 2409 ATGAACACACTCTCCATGGATTTGAACATATG--ACTTATTTGTAGGGGAGAGTCTCTG 2466

DB 1981 ATGAACACACTCTCCATGGGATTTGAACATATGAAGTTATTTGTAGGGGAGAGTCTCTG 2040  
QY 2467 AGGGGCAACACACAGAACACAGTCCCTCAGCCACAGCAGTCTCTTTTGTGTATCCA 2526  
DB 2041 AGGGGCAACACACAGAACACAGTCCCTCAGCCACAGCAGTCTCTTTTGTGTATCCA 2100  
QY 2527 CCCCCTCTTACCCTTTATCAGAGTGTGGCTGTGGTCTCTGTTGCTGCTGCTGCTGCTGCTG 2586  
DB 2101 CCCCCTCTTACCCTTTATCAGAGTGTGGCTGTGGTCTCTGTTGCTGCTGCTGCTGCTGCTG 2160  
QY 2587 CACAGGCAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2646  
DB 2161 CACAGGCAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220  
QY 2647 CTTTCTCTGTTGGTGTCTAATATTTGGTGTGGGTAGGATGAGGATGAGGATGAGGATGAGG 2706  
DB 2221 CTTTCTCTGTTGGTGTCTAATATTTGGTGTGGGTAGGATGAGGATGAGGATGAGGATGAGG 2280  
QY 2707 CTGAGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2766  
DB 2281 CTGAGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340  
QY 2767 CCCCAGATGCTTAACCCAGGACCTTGGAATTTCTACTCATCCCAAAATGATAATTCACAA 2826  
DB 2341 CCCCAGATGCTTAACCCAGGACCTTGGAATTTCTACTCATCCCAAAATGATAATTCACAA 2400  
QY 2827 TGCTGTTTACCAGGTTAGGTTGTGAAGAGGTAGAGGTAGAGGTAGAGGTAGAGGTAGAGGTAG 2886  
DB 2401 TGCTGTTTACCAGGTTAGGTTGTGAAGAGGTAGAGGTAGAGGTAGAGGTAGAGGTAGAGGT 2460  
QY 2887 GSGTTCCTTAACACACCTCTCTCTGTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2946  
DB 2461 GSGTTCCTTAACACACCTCTCTCTGTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
QY 2947 TCTACTCTCTTAGGACTGGGCTGATGAAGGCACTGCCCCAAATTTCCCTTACCCCAAC 3006  
DB 2521 TCTACTCTCTTAGGACTGGGCTGATGAAGGCACTGCCCCAAATTTCCCTTACCCCAAC 2580  
QY 3007 TTTCCCTTACCCCAACTTTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3066  
DB 2581 TTTCCCTTACCCCAACTTTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640  
QY 3067 CCAGAAGCACAAAGTGGGTTTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3126  
DB 2641 CCAGAAGCACAAAGTGGGTTTCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
QY 3127 GTGCTTGGGGAACTCTCACACAGAACTCAGGAGCACCCCTGCTGCTGCTGCTGCTGCTGCTGCT 3186  
DB 2701 GTGCTTGGGGAACTCTCACACAGAACTCAGGAGCACCCCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
QY 3187 TTATCTCTCAGGGGGGTTTAAAGTCCGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3246  
DB 2761 TTATCTCTCAGGGGGGTTTAAAGTCCGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2820  
QY 3247 GGTGAATTTTATCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3306  
DB 2821 GGTGAATTTTATCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2880  
QY 3307 AAGCTTCTTATATGTTTAAAAA 3330  
DB 2881 AAGCTTCTTATATGTTTAAAAA 2904

## RESULT 20

US-09-895-793-703  
; Sequence 703, Application US/09895793  
; Publication No. US20020192763A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.



```
QY 2109 GCTGGGGCTGGGGGCTCCCTCTCTCTCTCCCGAGTCTTAGGGCTGCCCTGACTGGAGG 2168
Db 1681 GCTGGGGCTGGGGGCTCCCTCTCTCTCTCCCGAGTCTTAGGGCTGCCCTGACTGGAGG 1740
QY 2169 CCTTCCAAAGGGGTTTTCAGTCTGGACTTATACAGGGAGGCCAGAGGGCTCCATGCACGTG 2228
Db 1741 CCTTCCAAAGGGGTTTTCAGTCTGGACTTATACAGGGAGGCCAGAGGGCTCCATGCACGTG 1800
QY 2229 GAATCGGGGACTCTGAGGTGGATTACCCAGGCTCAGGGTTAAACAGTAGCCCTCTAGT 2288
Db 1801 GAATCGGGGACTCTGAGGTGGATTACCCAGGCTCAGGGTTAAACAGTAGCCCTCTAGT 1860
QY 2289 TGAGACACACTAGAGAGGGTTTTTGGGAGCTGAATAAATCAGTCACTGGTTTCCCA 2348
Db 1861 TGAGACACACTAGAGAGGGTTTTTGGGAGCTGAATAAATCAGTCACTGGTTTCCCA 1920
QY 2349 TCTCTAAGCCCTTAACTGAGCTGAGCTTCGTTTAAATAGTCTCTGATGGGAGTTCTAGG 2408
Db 1921 TCTCTAAGCCCTTAACTGAGCTGAGCTTCGTTTAAATAGTCTCTGATGGGAGTTCTAGG 1980
QY 2409 ATGAACACTCTCTCATGGGATTGAACATATG--ACTTATTGTAGGGAGAGTCTCTG 2466
Db 1981 ATGAACACTCTCTCATGGGATTGAACATATGAAAGTTATTGTAGGGAGAGTCTCTG 2040
QY 2467 AGGGGACACACACAGAACCCAGTCCCTCAGCCACAGCACTGCTCTTTTGTGATCCCA 2526
Db 2041 AGGGGACACACACAGAACCCAGTCCCTCAGCCACAGCACTGCTCTTTTGTGATCCCA 2100
QY 2527 CCCCCCTCTTACCTTTATCAGGATGTGGCTGTGGTCTTCTCTGTCCTATCAGAGAGA 2586
Db 2101 CCCCCCTCTTACCTTTATCAGGATGTGGCTGTGGTCTTCTCTGTCCTATCAGAGAGA 2160
QY 2587 CACAGCACTTAAATTAATTAATTAATTAATTAACAAAGTAGAGGAATCCATTGCTAG 2646
Db 2161 CACAGCACTTAAATTAATTAATTAATTAATTAACAAAGTAGAGGAATCCATTGCTAG 2220
QY 2647 CTTTCTGTGTGGTCTTAATTAATTTGGTAGGGTGGGGATCCCAACAATCAGTCTCC 2706
Db 2221 CTTTCTGTGTGGTCTTAATTAATTTGGTAGGGTGGGGATCCCAACAATCAGTCTCC 2280
QY 2707 CTGAGATAGTGTGCTATGGGCTGATCATTTGCCAGAACTCTTCTCTCGGGGCTGCGC 2766
Db 2281 CTGAGATAGTGTGCTATGGGCTGATCATTTGCCAGAACTCTTCTCTCGGGGCTGCGC 2340
QY 2767 CCCCCAAATGCCCTAACCCAGGACCTTGGAAATTTCTACTCATCCCAAAATGATAATCCAAA 2826
Db 2341 CCCCCAAATGCCCTAACCCAGGACCTTGGAAATTTCTACTCATCCCAAAATGATAATCCAAA 2400
QY 2827 TGCTGTATCCCAAGGTTAGGGTGTGAAGGAAGGTAGAGGTGGGGCTTCAGGTCTCAAC 2886
Db 2401 TGCTGTATCCCAAGGTTAGGGTGTGAAGGAAGGTAGAGGTGGGGCTTCAGGTCTCAAC 2460
QY 2887 GGCTTCCCAACACCCCTCTCTCTCTTGTGCCCCAGGCTGTTCCCGCCACTTCCACTCCCG 2946
Db 2461 GGCTTCCCAACACCCCTCTCTCTCTTGTGCCCCAGGCTGTTCCCGCCACTTCCACTCCCG 2520
QY 2947 TCTACTCTCTTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCGCTACCCCGAAC 3006
Db 2521 TCTACTCTCTTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCGCTACCCCGAAC 2580
QY 3007 TTTTCCCGCTACCCCGAACTTTCCCGACAGCTCCACAAACCTGTGTTGGAGCTACTGCAGGA 3066
Db 2581 TTTTCCCGCTACCCCGAACTTTCCCGACAGCTCCACAAACCTGTGTTGGAGCTACTGCAGGA 2640
QY 3067 CCAGAGCAAAAGTGGGTTTCCCAAGCTTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
Db 2641 CCAGAGCAAAAGTGGGTTTCCCAAGCTTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700
QY 3127 GTGCTTGGGAAATCTCACAGAAACTCAGGACACCCCTGCCTGAGCTAAGGGAGGTC 3186
Db 2701 GTGCTTGGGAAATCTCACAGAAACTCAGGACACCCCTGCCTGAGCTAAGGGAGGTC 2760
```

3187 TTATCTCTCAGGGGGTTTAAAGTCGCTTGTGCAATAATGTCGCTTATTTATTAGCGG 3246  
Db 2761 TTATCTCTCAGGGGGTTTAAAGTCGCTTGTGCAATAATGTCGCTTATTTATTAGCGG 2820  
QY 3247 GGTGAATATTTTACTTGTAAAGTGAGCAATCAGAGTATAATGTTTATGTCGACAAAATTA 3306  
Db 2821 GGTGAATATTTTACTTGTAAAGTGAGCAATCAGAGTATAATGTTTATGTCGACAAAATTA 2880  
QY 3307 AAGGCTTCTTATATGTTTAAAAA 3330  
Db 2881 AAGGCTTCTTATATGTTTAAAAA 2904

RESULT 21  
US-09-895-814-703  
; Sequence 703, Application US/09895814  
; Publication No. US20020193296A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darlick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C26  
; CURRENT APPLICATION NUMBER: US/09/895.814  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 990  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 703  
; LENGTH: 2904  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-895-814-703

Query Match 64.4%; Score 2196.4; DB 9; Length 2904;  
Best Local Similarity 88.2%; Pred. No. 0;  
Matches 2560; Conservative 0; Mismatches 1; Indels 343; Gaps 2;

QY 770 GTCATGCTTCATGATCAGTCTGGGGCTGCTGGGCTACCTCTCGCTGCCATTGAC 829  
Db 1 GTCATGCTTCATGATCAGTCTGGGGCTGCTGGGCTACCTCTCGCTGCCATTGAC 60  
QY 830 TGGGACACAGTGCCTTGGGGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCTG 889  
Db 61 TGGGACACAGTGCCTTGGGGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCTG 120  
QY 890 CTCACCTCTATCTTCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 949  
Db 121 CTCACCTCTATCTTCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 950 CTGGGGCCCCCAGGAGGAGGAGGCTGCTGGGGCCCCCTCTCTCTGCTGCCCTGCTGCT 1009  
Db 181 CTGGGGCCCCCAGGAGGAGGAGGCTGCTGGGGCCCCCTCTCTCTGCTGCCCTGCTGCT 240

QY 1010 CCATGCGGGCGCGCTTGGCTTTCCGAAACCTGGGGCGCCCTGCTTCCCGGCTGCACCAG 1069  
DB 241 CCATGCGGGCGCGCTTGGCTTTCCGAAACCTGGGGCGCCCTGCTTCCCGGCTGCACCAG 300  
QY 1070 CTGTGCTGCCGATGCCCGACCCCTGCGCGGCTTTCGTGGCTGAGCTGTGCAGCTGG 1129  
DB 301 CTGTGCTGCCGATGCCCGACCCCTGCGCGGCTTTCGTGGCTGAGCTGTGCAGCTGG 360  
QY 1130 ATGGCACTCATGACCTTACGCTGTGTTTACAGGGATTTTCGTGGCGAGGGGCTGTACCAG 1189  
DB 361 ATGGCACTCATGACCTTACGCTGTGTTTACAGGGATTTTCGTGGCGAGGGGCTGTACCAG 420  
QY 1190 GGGCTGCCAGAGCTGAGCCGGGACCGAGGCCGCGAGACATATGATGAA 1240  
DB 421 GGGCTGCCAGAGCTGAGCCGGGACCGAGGCCGCGAGACATATGATGAAAGAGGCT 480  
QY 1241 ----- 1240  
DB 481 CTGGCTGCTTAGGAGTCTGATCAGATGCTGTTGCCCCAGTTTGACAGAAAGAGCGGA 540  
QY 1241 ----- 1240  
DB 541 GCTTATTCAAAGTCTAGAGGAGTGGAGAGTTAAGGCTGGATTTCAGATCTGCTGCTT 600  
QY 1241 ----- 1240  
DB 601 CCAGCGCAGTGTGCCCTCTGCTCCCAACGACTTTCCAAATATCTACACAGCGCCTT 660  
QY 1241 ----- 1240  
DB 661 CCAGCTCAGCGCTCTAGAGCGCTTGAAGCCTATGCCAGCTCTTTGTTGTTCCCTC 720  
QY 1241 ----- 1240  
DB 721 TCACCCGCTGTCTCACAGCTGAGACTCCAGGAAACCTTCAGACTACCTTCTCTGCC 780  
QY 1241 ----- 1268  
DB 781 TTCAGCAAGGGCGTTGCCACATCTCTGAGGGCTTCGGATGGCAGCTGGGGTGT 840  
QY 1269 TCCTGAGTGGCCATCTCCCTGGTCTCTCTGTCATGACCGGCTGTGTCAGGGAT 1328  
DB 841 TCCTGAGTGGCCATCTCCCTGGTCTCTCTGTCATGACCGGCTGTGTCAGGGAT 900  
QY 1329 TCGGCACCTGAGCAGCTATTGGCCAGTGTGGCAGCTTCCCTGTGGCTGCCGGTGCCA 1388  
DB 901 TCGGCACCTGAGCAGCTATTGGCCAGTGTGGCAGCTTCCCTGTGGCTGCCGGTGCCA 960  
QY 1389 CATGCTGTCCACAGTGTGGCGCTGTGTGACAGCTTTCAGCGCGCTCACCGGGTTCACCT 1448  
DB 961 CATGCTGTCCACAGTGTGGCGCTGTGTGACAGCTTTCAGCGCGCTCACCGGGTTCACCT 1020  
QY 1449 TCTACGCTGTGAGATCTGCCCTACACACTGGGCTCCCTTACACCGGGAGAGCAGG 1508  
DB 1021 TCTACGCTGTGAGATCTGCCCTACACACTGGGCTCCCTTACACCGGGAGAGCAGG 1080  
QY 1509 TGTTCCTGCCAAATACCGAGGGGACACTGGAGTGTGTAGCAGTACGACAGCCTGATGA 1568  
DB 1081 TGTTCCTGCCAAATACCGAGGGGACACTGGAGTGTGTAGCAGTACGACAGCCTGATGA 1140  
QY 1569 CCAGCTTCTGCCAGGCGCTTAAGCCTGTGAGCTCCCTTCCCTTAATGGACAGCTGGGTGCTG 1628  
DB 1141 CCAGCTTCTGCCAGGCGCTTAAGCCTGTGAGCTCCCTTCCCTTAATGGACAGCTGGGTGCTG 1200  
QY 1629 GAGCAGTGGGCTGTGCCACCTCCACCGGCTGTGCGGGGCTCTGCGCTGTGATGCT 1688  
DB 1201 GAGCAGTGGGCTGTGCCACCTCCACCGGCTGTGCGGGGCTCTGCGCTGTGATGCT 1260  
QY 1689 CCGTACGTGTGTGGTGGGTCAGCCACCGAGGCCAGGCTGTGTCGCGGGGCTGTGATCT 1748  
DB 1261 CCGTACGTGTGTGGTGGGTCAGCCACCGAGGCCAGGCTGTGTCGCGGGGCTGTGATCT 1320  
QY 1749 GCCTGGACCTCGGCATCTCGATAGTGCCTTCTGCTGTCCACAGGTGGCGCCCTCCTGT 1808

DB 1321 GCCTGGACCTCGGCATCTCTGGATAGTGCCTTCTCTGCTGCCAGTGGCCCCCTCCTGT 1380  
QY 1809 TTATGGGCTCAATTTGTCAGCTCAGCAGTCTGTCACTGCTATATGTTCTGTGCCGAG 1868  
DB 1381 TTATGGGCTCAATTTGTCAGCTCAGCAGTCTGTCACTGCTATATGTTCTGTGCCGAG 1440  
QY 1869 GCCTGGGCTGTGGTGGCGCATTTACTTTGCTACACAGTGTAGTATTGACAAAGGAGCTGG 1928  
DB 1441 GCCTGGGCTGTGGTGGCGCATTTACTTTGCTACACAGTGTAGTATTGACAAAGGAGCTGG 1500  
QY 1929 CCAATATCTCAGCGTAGAAAACCTTCCAGCACATTTGGGTGGAGGCTGCTCCTCAGTGGT 1988  
DB 1501 CCAATATCTCAGCGTAGAAAACCTTCCAGCACATTTGGGTGGAGGCTGCTCCTCAGTGGT 1560  
QY 1989 CCAGCTTCCCGGCTCTGTTAGCCCATGGGGCTGCGGGCTGGCGGCCAGTTTCTGTG 2048  
DB 1561 CCCAGCTTCCCGGCTCTGTTAGCCCATGGGGCTGCGGGCTGGCGGCCAGTTTCTGTG 1620  
QY 2049 CTGCAAGTAATGTGGCTCTCTGCTGCCACCTGCTGCTCAGGTGCGTAGCTGCACA 2108  
DB 1621 CTGCAAGTAATGTGGCTCTCTGCTGCCACCTGCTGCTCAGGTGCGTAGCTGCACA 1680  
QY 2109 GCTGGGGCTGTGGGGCTTCCCTCTCTCTCTCCAGTCTCTAGGGCTGCTGACTGGAG 2168  
DB 1681 GCTGGGGCTGTGGGGCTTCCCTCTCTCTCTCTCCAGTCTCTAGGGCTGCTGACTGGAG 1740  
QY 2169 CCTTCCAGGGGCTTTCAGTCTGGACTTATACAGGAGGCCAGAGGCTCCATGCACTG 2228  
DB 1741 CCTTCCAGGGGCTTTCAGTCTGGACTTATACAGGAGGCCAGAGGCTTCCATGCACTG 1800  
QY 2229 GAATGCGGGGACTCTGAGGTGATTAACCCAGGCTCAGGGTTAACAGCTAGCCTCTAGT 2288  
DB 1801 GAATGCGGGGACTCTGAGGTGATTAACCCAGGCTCAGGGTTAACAGCTAGCCTCTAGT 1860  
QY 2289 TGAGACACCTAGAGAGGGTTTTTGGGAGCTGAATAACTCAGTCACTGGTTTCCCA 2348  
DB 1861 TGAGACACCTAGAGAGGGTTTTTGGGAGCTGAATAACTCAGTCACTGGTTTCCCA 1920  
QY 2349 TCTCTAAGCCCTTAACCTCAGCTTCTGTTTAAATAGTCTTTCATGGAGTTTCTAGG 2408  
DB 1921 TCTCTAAGCCCTTAACCTCAGCTTCTGTTTAAATAGTCTTTCATGGAGTTTCTAGG 1980  
QY 2409 ATGAACACTCTCCATGGGATTTGAACATATG - ACTTATTTGTAGGGAAGAGTCTG 2466  
DB 1981 ATGAACACTCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGAAGAGTCTG 2040  
QY 2467 AGGGCAACACACAGAACAGCTCCCTCAGCCACAGCACTGCTTTTGTGATCCA 2526  
DB 2041 AGGGCAACACACAGAACAGCTCCCTCAGCCACAGCACTGCTTTTGTGATCCA 2100  
QY 2527 CCCCCCTTACCTTTTATCAGGATGTGGCTGTTGGTCTCTTCTGTTGCCATCACAGAGA 2586  
DB 2101 CCCCCCTTACCTTTTATCAGGATGTGGCTGTTGGTCTCTTCTGTTGCCATCACAGAGA 2160  
QY 2587 CACAGCATTTAAATATTTAACTTATTTAAACAAAGTAGAAGGAATCCATTGCTAG 2646  
DB 2161 CACAGCATTTAAATATTTAACTTATTTAAACAAAGTAGAAGGAATCCATTGCTAG 2220  
QY 2647 CTTTCTGTGTGGTCTTAATATTTGGTGGGGATCCCCAACAAATCAGTCCC 2706  
DB 2221 CTTTCTGTGTGGTCTTAATATTTGGTGGGGATCCCCAACAAATCAGTCCC 2280  
QY 2707 CTGAGATAGCTGGTCAATTTGGGCTGATCATTTGCCAGAAATCTTCTCTCTGGGGCTGGCC 2766  
DB 2281 CTGAGATAGCTGGTCAATTTGGGCTGATCATTTGCCAGAAATCTTCTCTCTGGGGCTGGCC 2340  
QY 2767 CCCCCAAATGCTTAACCCAGGACCTTGGAAATTTACTCATCCCCAAATGATATTTCCAA 2826  
DB 2341 CCCCCAAATGCTTAACCCAGGACCTTGGAAATTTACTCATCCCCAAATGATATTTCCAA 2400  
QY 2827 TGTGTTACCCCAAGGTTAGGGTGTGAAGAGGTTAGAGGGTGGGGCTTCAGTCTCAAC 2886



1509 TGTTCTCCGCAATACCGAGGGACACTGAGGTGCTAGCAGTACAGGACAGCTGATGA 1568  
1081 TGTTCTCCGCAATACCGAGGGACACTGAGGTGCTAGCAGTACAGGACAGCTGATGA 1140  
1569 CCAGCTTCTGCGAGGCTTACGCTGAGCTCCCTTCCCTTAATGAGACAGTGGTGGTCTG 1628  
1141 CCAGCTTCTGCGAGGCTTACGCTGAGCTCCCTTCCCTTAATGAGACAGTGGTGGTCTG 1200  
1629 GAGGAGTGGCTGCTCCACCTCCACCGGCTCTCGGGGCTCTGCTGCTGATGCTCT 1688  
1201 GAGGAGTGGCTGCTCCACCTCCACCGGCTCTCGGGGCTCTGCTGCTGATGCTCT 1260  
1689 CCGTACGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1748  
1261 CCGTACGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1320  
1749 GCCTGAGCTCGCATCTGATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1808  
1321 GCCTGAGCTCGCATCTGATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
1809 TTATGGGCTCCATTTGCTGAGCTCAGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1928  
1381 TTATGGGCTCCATTTGCTGAGCTCAGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
1869 SCCTGGGCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1928  
1441 GCCTGGGCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1500  
1929 CCAATACTACGCTAGCAAACTTCCAGCAGTGGGCTGGAGGCTGCTGCTGCTGCTGCTGCT 1988  
1501 CCAATACTACGCTAGCAAACTTCCAGCAGTGGGCTGGAGGCTGCTGCTGCTGCTGCTGCT 1560  
1989 CCCAGCTCCCGCTCTGTTAGCCCTATGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 2048  
1561 CCCAGCTCCCGCTCTGTTAGCCCTATGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1620  
2049 CTGCAAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2108  
1621 CTGCAAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
2109 GCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 2168  
1681 GCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1740  
2169 CCTTCAAGGGGTTTACGCTGACCTTATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2228  
1741 CCTTCAAGGGGTTTACGCTGACCTTATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800  
2229 CAATCGGGGACTCTGAGGTGGATTACCGAGGCTCAGGCTTAAACAGCTAGCTCCTCTAGT 2288  
1801 GAATCGGGGACTCTGAGGTGGATTACCGAGGCTCAGGCTTAAACAGCTAGCTCCTCTAGT 1860  
2289 TGAGACACCTAGAGAGGGTTTTGGGAGCTGAATAACTCAGTACCTGCTGCTGCTGCTGCTG 2348  
1861 TGAGACACCTAGAGAGGGTTTTGGGAGCTGAATAACTCAGTACCTGCTGCTGCTGCTGCTG 1920  
2349 TCTCTAAGCCCTTAACTGAGCTTCTGTTAATGAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 2408  
1921 TCTCTAAGCCCTTAACTGAGCTTCTGTTAATGAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1980  
2409 ATGAACACTCTCTCCATGGGATTTGAACATATG - ACTTATTTTGTAGGGGAAGAGTCCCTG 2466  
1981 ATGAACACTCTCTCCATGGGATTTGAACATATGAAAGTATTTGTAGGGGAAGAGTCCCTG 2040  
2467 AGGGCAACACACAAAGAACAGGTCCTCAGCCACAGCACTGCTGCTGCTGCTGCTGCTGCTG 2526  
2041 AGGGCAACACACAAAGAACAGGTCCTCAGCCACAGCACTGCTGCTGCTGCTGCTGCTGCTG 2100  
2527 CCCCCCTTACCTTTTATCAGGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2586  
2101 CCCCCCTTACCTTTTATCAGGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
2587 CACAGCATTTAAATATTTAACTTATTTATTAACAAAGTAGAAGGGAATCCATTCCTAG 2646

2161 CACAGGCATTAAATATTTAACTTATTTAAACAAGTAGAAGGAATCCATTCCTAG 2220  
2647 CTTTTCTGTTGGTCTTAATATTTGGTAGGGTGGGGATCCCAACAATCAGGTCC 2706  
2221 CTTTTCTGTTGGTCTTAATATTTGGTAGGGTGGGGATCCCAACAATCAGGTCC 2280  
2707 CTGAGATAGTGTGCTATTTGGGCTGATCATTGCCAGAACTTCTTCTCTGCTGGGCTGCGC 2766  
2281 CTGAGATAGTGTGCTATTTGGGCTGATCATTGCCAGAACTTCTTCTCTGCTGGGCTGCGC 2340  
2767 CCCCCAAATGCTTAACCCAGGACCTTTGAAATTTCTACTCATCCCAATGATATTTCCAAA 2826  
2341 CCCCCAAATGCTTAACCCAGGACCTTTGAAATTTCTACTCATCCCAATGATATTTCCAAA 2400  
2827 TGGTGTACCCAGGCTAGGCTGTTGAAGAAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886  
2401 TGGTGTACCCAGGCTAGGCTGTTGAAGAAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460  
2887 GGCTTCCCTAACCCAGGCTTCTTCTGCTGGCCAGGCTGCTTCCCTCCCTTCCCTCCCT 2946  
2461 GGCTTCCCTAACCCAGGCTTCTTCTGCTGGCCAGGCTGCTTCCCTCCCTTCCCTCCCT 2520  
2947 TCTACTCTCTAGGACTGGGCTGATGAAGGACCTGCCCAAAATTTCCCTTACCCTCCAAAC 3006  
2521 TCTACTCTCTAGGACTGGGCTGATGAAGGACCTGCCCAAAATTTCCCTTACCCTCCAAAC 2580  
3007 TTTCCCTTACCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 3066  
2581 TTTCCCTTACCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 2640  
3067 CCAGAAGCACAAAGTGGGTTTTCCCAAGCCTTTGTCATCTCAGCCCCCAGAGTATCT 3126  
2641 CCAGAAGCACAAAGTGGGTTTTCCCAAGCCTTTGTCATCTCAGCCCCCAGAGTATCT 2700  
3127 GTGCTTGGGAACTCTCACAGAACTCAGGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3186  
2701 GTGCTTGGGAACTCTCACAGAACTCAGGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
3187 TTATCTCTCAGGGGGGTTTTAAGTGGGTTTTGCAATAATGCTGCTTATTTATTTAGCGG 3246  
2761 TTATCTCTCAGGGGGGTTTTAAGTGGGTTTTGCAATAATGCTGCTTATTTATTTAGCGG 2820  
3247 GGTGATATTTTATCTGTAAGTGAGCAATCAGAGTATTAATGTTTATGTCACAAATTA 3306  
2821 GGTGATATTTTATCTGTAAGTGAGCAATCAGAGTATTAATGTTTATGTCACAAATTA 2880  
3307 AAGGCTTCTTATATGTTTTAAAAA 3330  
2881 AAGGCTTCTTATATGTTTTAAAAA 2904

## RESULT 23

US-09-780-669-703  
; Sequence 703, Application US/09780669  
; Patent No. US20020051977A1

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqul  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.







```
Db 1861 TGAGACACACCTTAGAGAGAGGTTTTGGAGCTGAATAACTCAGTCACCTGGGTTCCCA 1920
QY TCTTAAGCCCTTAACCTGCAGCTTCGTTAATAGTCTCTTGATGGAGTTCTAGG 2408
Db 1921 TCTTAAGCCCTTAACCTGCAGCTTCGTTAATAGTCTCTTGATGGAGTTCTAGG 1980
QY ATGAACACTCTCCATGGATTGAACATATG---ACTATTGTAGGGGAAGAGTCTG 2466
Db 1981 ATGAACACTCTCCATGGATTGAACATATGAAAGTTATTGTAGGGGAAGAGTCTG 2040
QY AGGGCAACACAGAACAGCTCCCTCAGCCACAGACTGCTTTTTGCTGATCCA 2526
Db 2041 AGGGCAACACAGAACAGCTCCCTCAGCCACAGACTGCTTTTTGCTGATCCA 2100
QY CCCCTCTTACCTTTATCAGGATGTGGCTGTGGTCTCTGTTGCCATCACAGAGA 2586
Db 2101 CCCCTCTTACCTTTATCAGGATGTGGCTGTGGTCTCTGTTGCCATCACAGAGA 2160
QY CACAGGCATTTAAATATTAACTTATTATTAACAAAGTAGAGGGGAATCCATTGCTAG 2646
Db 2161 CACAGGCATTTAAATATTAACTTATTATTAACAAAGTAGAGGGGAATCCATTGCTAG 2220
QY CTTTCTGTGTGGTCTTAATATTGGTAGGTGGGGATCCCAACATCAGGTCCC 2706
Db 2221 CTTTCTGTGTGGTCTTAATATTGGTAGGTGGGGATCCCAACATCAGGTCCC 2280
QY CTGAGATAGCTGGTCAATGGGCTGATCATTCGCAAGATCTTCTCTCTGGGCTCTGGCC 2766
Db 2281 CTGAGATAGCTGGTCAATGGGCTGATCATTCGCAAGATCTTCTCTCTGGGCTCTGGCC 2340
QY CCCCRAAATGCTTAACCCAGAGCTTGGAAATTTCTACTCATCCAAATGATAATTCAAA 2826
Db 2341 CCCCRAAATGCTTAACCCAGAGCTTGGAAATTTCTACTCATCCAAATGATAATTCAAA 2400
QY TGCTGTACCCAAAGTTAGGTTGTTGAAGAGGTAGAGGTGGGCTTCAAGTCTCAAC 2886
Db 2401 TGCTGTACCCAAAGTTAGGTTGTTGAAGAGGTAGAGGTGGGCTTCAAGTCTCAAC 2460
QY GGCCTTCCCTAACCCACCTCTCTCTTGGCCACAGCTGGTTCCCTCCCTCCCTCCCT 2946
Db 2461 GGCCTTCCCTAACCCACCTCTCTCTTGGCCACAGCTGGTTCCCTCCCTCCCTCCCT 2520
QY TCTACTCTCTTAGGACTGGGCTGATGAAGCACTGCCCAAAATTCCTCCCTACCCCAAC 3006
Db 2521 TCTACTCTCTTAGGACTGGGCTGATGAAGCACTGCCCAAAATTCCTCCCTACCCCAAC 2580
QY TTTCCCTTACCCCAACTTTCCCAACAGCTCCCAACCTGTTTGGAGCTACTGCAAGA 3066
Db 2581 TTTCCCTTACCCCAACTTTCCCAACAGCTCCCAACCTGTTTGGAGCTACTGCAAGA 2640
QY CCAGAAGCACAAAGTGGGTTTCCCAAGCTTTGTCATCTACGCCCCAGAGTATATCT 3126
Db 2641 CCAGAAGCACAAAGTGGGTTTCCCAAGCTTTGTCATCTACGCCCCAGAGTATATCT 2700
QY GTGCTTGGGGAATCTCACAGAAACTCAGGACACCTCCCTGCTGAGCTAAGGGAGTC 3186
Db 2701 GTGCTTGGGGAATCTCACAGAAACTCAGGACACCTCCCTGCTGAGCTAAGGGAGTC 2760
QY TTATCTCTCAGGGGGGTTTAAAGTCCCTTAAATATGCTCTTATTATTAGCGG 3246
Db 2761 TTATCTCTCAGGGGGGTTTAAAGTCCCTTAAATATGCTCTTATTATTAGCGG 2820
QY GGTGAATATTTTACTCTAAGTAGGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
Db 2821 GGTGAATATTTTACTCTAAGTAGGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880
QY AAGGCTTCTTATATGTTTAAAAA 3330
Db 2881 AAGGCTTCTTATATGTTTAAAAA 2904
```

RESULT 24

```
US-09-822-827-703
; Sequence 703, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 703
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-822-827-703
```

```
Query Match 64.4%; Score 2196.4; DB 10; Length 2904;
Best Local Similarity 88.2%; Pred. No. 0;
Matches 2560; Conservative 0; Mismatches 1; Indels 343; Gaps 2;

QY 770 GTCTATGCCCTTCATGATCAGTCTTGGGGCTGCTGGCTACCTCCTGCTGCTGCTGCTGAC 829
Db 1 GTCTATGCCCTTCATGATCAGTCTTGGGGCTGCTGGCTACCTCCTGCTGCTGCTGCTGAC 60
QY 830 TGGGACACCAAGTGGCTTGGGGCTTACCTGCGGACCCAGAGGAGTGCCTCTTTGGCCCTG 889
Db 61 TGGGACACCAAGTGGCTTGGGGCTTACCTGCGGACCCAGAGGAGTGCCTCTTTGGCCCTG 120
QY 890 CTCACTCTCATCTTCTCAGTCTAGTACGACGACACTGCTGTGCTGCTGAGGAGGAGCG 949
Db 121 CTCACTCTCATCTTCTCAGTCTAGTACGACGACACTGCTGTGCTGCTGAGGAGGAGCG 180
QY 950 CTGGGCCCCACCGAGCAGCAAGGCTGCTGGGGCCCCCTCTCTTGTGCCCCACCTGCTCT 1009
Db 181 CTGGGCCCCACCGAGCAGCAAGGCTGCTGGGGCCCCCTCTCTTGTGCCCCACCTGCTCT 240
QY 1010 CCATGGCGGGCGCTTGGCTTTCCGGAACCTGGGGCGCTGCTTCCCGGGTGCACACAG 1069
Db 241 CCATGGCGGGCGCTTGGCTTTCCGGAACCTGGGGCGCTGCTTCCCGGGTGCACACAG 300
QY 1070 CTGTGCTGGCGATGCCCGCACCTTGGCGGGCTTCTGTGCTGCTGCTGCTGCTGCTG 1129
Db 301 CTGTGCTGGCGATGCCCGCACCTTGGCGGGCTTCTGTGCTGCTGCTGCTGCTGCTGCTG 360
QY 1130 ATGGCACTCATGACCTTACGCTGTTTACACGATTTTACGCGGAGGCTGTGCTGCTGCTG 1189
Db 361 ATGGCACTCATGACCTTACGCTGTTTACACGATTTTACGCGGAGGCTGTGCTGCTGCTG 420
QY 1190 GGCGTGGCCAGCTGAGCGGGGACCGGAGGCGGCGGAGACTATGATGAA 1240
Db 421 GGCGTGGCCAGCTGAGCGGGGACCGGAGGCGGCGGAGACTATGATGAAAGAGGCT 480
QY 1241 1241
Db 481 CTGGCTGCTTAGAGTCTGATCAGAGTCTGTTGCCAGTTTGCACAGAAAGGCGGA 540
QY 1241 1241
Db 541 GCTTATTCAAAGTCTAGAGGAGTGGAGGTTAAGGCTTGAATTTTTCAGATCTGCTGCTT 600
QY 1241 1241
Db 601 CCAGCGCAGTGTGCCCTCTGTCTCCCAACGACTTTTCCAAATTAATCTCACCAGGCGCT 660
QY 1241 1241
Db 661 CCAGCTCAGGCTCTTAGAAGGCTTGAAGGCTATGGCCAGTGTCTTTGTGCTTCCCTC 720
QY 1241 1241
Db 721 TCACCGGCTGTCTCAGACTCAGACTCCAGGAAACCTTCCAGACTACCTTCTCTCTGCC 780
```

QY	1241	-----GGCGTTCCGATGGCAGCCTGGGCGTGT	1268
Db	781	TTTCAGCAAGGGCGTTGCCACATCTCTGAGGGCGTTCGGATGGCAGCCTGGGCGTGT	840
QY	1269	TGCTGGAGTGGCCATCTCCCTGGTGCTTCTCTCTGGTCATGACCGGCTGGTGGCAGCGAT	1328
Db	841	TCTGCGAGTGGCCATCTCCCTGGTGCTTCTCTCTGGTCATGACCGGCTGGTGGCAGCGAT	900
QY	1329	TGGCACTCGACGAGTCTATTTTGGCCAGTGTGGCAGCTTCCCTGTGGCTGCCGTGGCCA	1388
Db	901	TGGCACTCGACGAGTCTATTTTGGCCAGTGTGGCAGCTTCCCTGTGGCTGCCGTGGCCA	960
QY	1389	CATGCTGTCCACAGTGTGGCCGTGGTGACAGCTTTCAGCCGCCCTCACCGGGTTCACT	1448
Db	961	CATGCTGTCCACAGTGTGGCCGTGGTGACAGCTTTCAGCCGCCCTCACCGGGTTCACT	1020
QY	1449	TCTCAGCCCTGCAGATCCCTGCCTACACACTGSCCTCCCTCTACCAACGGGAGACAGG	1508
Db	1021	TCTCAGCCCTGCAGATCCCTGCCTACACACTGSCCTCCCTCTACCAACGGGAGACAGG	1080
QY	1509	TGTTCTTGCCCAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAAGACACGCTGATGA	1568
Db	1081	TGTTCTTGCCCAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAAGACACGCTGATGA	1140
QY	1569	CCAGCTTCTGCGAGGCCCTAAGCTGGAGTCCCTTCCCTAATGGACAGTGGGTGCTG	1628
Db	1141	CCAGCTTCTGCGAGGCCCTAAGCTGGAGTCCCTTCCCTAATGGACAGTGGGTGCTG	1200
QY	1629	GAGGAGTGGCTGCTCCACACTCCACCGCGCTCGCGGGGCTCTGCCTCTGATGCTCT	1688
Db	1201	GAGGAGTGGCTGCTCCACACTCCACCGCGCTCTGCAGGGGCTCTGCCTCTGATGCTCT	1260
QY	1689	CCGTAGCTGTGTGGGTGAGCCACCGAGGCCAGGGTGTTCGGGCCGGGGCATCT	1748
Db	1261	CCGTAGCTGTGTGGGTGAGCCACCGAGGCCAGGGTGTTCGGGCCGGGGCATCT	1320
QY	1749	GCCTGAGACCTCGGCATCTCGGATAGTGCCTTCTGCTGTCACAGTGGCCCCCATCCCTGT	1808
Db	1321	GCCTGAGACCTCGGCATCTCGGATAGTGCCTTCTGCTGTCACAGTGGCCCCCATCCCTGT	1380
QY	1809	TTATGGCTCCATGTGCCAGCTCAGCCAGTCTGTCACTGGCTATATGTGTCTGCCGCAG	1868
Db	1381	TTATGGCTCCATGTGCCAGCTCAGCCAGTCTGTCACTGGCTATATGTGTCTGCCGCAG	1440
QY	1869	GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGTACAAAGCGACTGG	1928
Db	1441	GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGTACAAAGCGACTGG	1500
QY	1929	CCAAATACTCAGCGTAGAAACCTTCACGACATTTGGGTGGAGGGCTTGCCTCACTGGGT	1988
Db	1501	CCAAATACTCAGCGTAGAAACCTTCACGACATTTGGGTGGAGGGCTTGCCTCACTGGGT	1560
QY	1989	CCCAGCTTCCCGCTCTCTGTAGCCCATGGGGCTGCCGGCTGGCCGCCAGTTTCTGTTG	2048
Db	1561	CCCAGCTTCCCGCTCTCTGTAGCCCATGGGGCTGCCGGCTGGCCGCCAGTTTCTGTTG	1620
QY	2049	CTGCCAAAGTAATGTGGCTCTGTGTGCCACCTGTGCTGCTGAGTGGCTAGCTGCACA	2108
Db	1621	CTGCCAAAGTAATGTGGCTCTGTGCTGCCACCTGTGCTGCTGAGTGGCTAGCTGCACA	1680
QY	2109	GCTGGGGCTGGGGCGTCCCTCTCTCTCCCACTCTAGGGCTGGCTGACTGAGG	2168
Db	1681	GCTGGGGCTGGGGCGTCCCTCTCTCTCCCACTCTAGGGCTGGCTGACTGAGG	1740
QY	2169	CCTTTCCAAGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAAGGCTTCCATGCACTG	2228
Db	1741	CCTTTCCAAGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAAGGCTTCCATGCACTG	1800
QY	2229	GAATGCGGGGACTCTGCAGTGGATTAACCAAGGCTCAGGGTTTACAGCTTAGCCTCTAGT	2288
Db	1801	GAATGCGGGGACTCTGCAGTGGATTAACCAAGGCTCAGGGTTTACAGCTTAGCCTCTAGT	1860

## RESULT 25

US-10-012-896-702  
; Sequence 702, Application US/10012896  
; Publication No. US20020183251A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNellil, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Wantanabe, Yoshihiro  
; APPLICANT: Meagher, Madeleine Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C27  
; CURRENT APPLICATION NUMBER: US/10/012,896  
; NUMBER OF SEQ ID NOS: 1011  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 702  
; LENGTH: 4894  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-012-896-702

Query Match 62.8%; Score 2142.8; DB 9; Length 4894;  
Best Local Similarity 80.1%; Pred. No. 0;  
Matches 2872; Conservative 1; Mismatches 3; Indels 708; Gaps 2;

QY	455	GCATTGCTCAGTCTGGGCTGGTGTGTGTCCTAGCTTCTAGGCTCAGCCAGTGACCAC	514
DB	1311	GCATTGCTCAGTCTGGGCTGGTGTGTGTCCTAGCTTCTAGGCTCAGCCAGTGACCAC	1370
QY	515	TGGCTGTGACGCTATGGCCGCGCGCTTCTATCTGGGCTGCTTGGGCTATCTG	574
DB	1371	TGGCTGTGACGCTATGGCCGCGCGCTTCTATCTGGGCTGCTTGGGCTATCTG	1430
QY	575	CTGAGCCCTTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGTGTGCGCCGATCCC	634
DB	1431	CTGAGCCCTTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGTGTGCGCCGATCCC	1490
QY	635	AGGCCCTTGGAGCTGGGCTGCTCATCTGGGCTGGGCTGGGCTGTGTGGCCAG	694
DB	1491	AGGCCCTTGGAGCTGGGCTGCTCATCTGGGCTGGGCTGGGCTGTGTGGCCAG	1550
QY	695	GTGTGCTTCACTGAGGCGCTGCTCTCTCATCTTCTGAGGCTGTGTGGGCTGCTG	754
DB	1551	GTGTGCTTCACTGAGGCGCTGCTCTCTCATCTTCTGAGGCTGTGTGGGCTGCTG	1610
QY	755	CGCCAGGCTTACTGTGTATGCTTCTATGATCATGCTTGGGGCTGCTGGGCTACCTC	814
DB	1611	CGCCAGGCTTACTGTGTATGCTTCTATGATCATGCTTGGGGCTGCTGGGCTACCTC	1670
QY	815	CTGCTGCTTACTGAGGCTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	874
DB	1671	CTGCTGCTTACTGAGGCTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1730

QY	875	TGCCTTTTGGCTGTCTACCCCTCATCTTCTACCTGTGCTAGCAGCACACTGTGGTG	934
DB	1731	TGCCTTTTGGCTGTCTACCCCTCATCTTCTACCTGTGCTAGCAGCACACTGTGGTG	1790
QY	935	GCTGAGGAGGAGGCTGGGCCCCACCGAGCCAGCAGAGGGCTGTGGGCCCTCTCTTG	994
DB	1791	GCTGAGGAGGAGGCTGGGCCCCACCGAGCCAGCAGAGGGCTGTGGGCCCTCTCTTG	1850
QY	995	TGCGCCCACTGCTGTCTCATGCCGGGCGCTTGGCTTTCGGGAACCTTGGGCGCTGCTT	1054
DB	1851	TGCGCCCACTGCTGTCTCATGCCGGGCGCTTGGCTTTCGGGAACCTTGGGCGCTGCTT	1910
QY	1055	CCCCGGCTGCACCACTGTGTCTGCCGATCCCGCACCTCCCGCACCTCCCGCACCT	1114
DB	1911	CCCCGGCTGCACCACTGTGTCTGCCGATCCCGCACCTCCCGCACCTCCCGCACCT	1970
QY	1115	GAGCTGTGACGCTGGATGGACATCATGACCTTACGCTGTGTTTACACGGATTTCTG	1174
DB	1971	GAGCTGTGACGCTGGATGGACATCATGACCTTACGCTGTGTTTACACGGATTTCTG	2030
QY	1175	GAGGGCTGTACCAAGGCGCTGCCAGAGCTGAGCGGGACCGGAGGCGGAGACACTAT	1234
DB	2031	GAGGGCTGTACCAAGGCGCTGCCAGAGCTGAGCGGGACCGGAGGCGGAGACACTAT	2090
QY	1235	GATCA-----	1239
DB	2091	GATGAAGGTAAGGCTTTGGCAGCAGCAGAGGCTGTGTGGGAGCGCCACCGAGACG	2150
QY	1240	-----	1239
DB	2151	ACACTGGGGCTGTGTCTGGGCTGGTGGCTTCTCCATCTTGGCCCCGACTTCTCTG	2210
QY	1240	-----	1239
DB	2211	AAAGTGGGATGGACCCCATCTGCATACAGGGCTTCTCATGGGTGTGGAACATCTCTG	2270
QY	1240	-----	1239
DB	2271	TGCGGTTTTCAGAAAGGCTTGGCTGTCTAGGAGTCTGATCAGAGCTGTGCCCCAGTT	2330
QY	1240	-----	1239
DB	2331	TGACAGAAGAAAGGCGGAGCTTATTCAAGTCTAGAGGAGTGGAGGAGTTAAGGCTGG	2390
QY	1240	-----	1239
DB	2391	ATTTGAGATCTGCTGGTTCCAGCGCAGTGTGCCCTCTGCTCCCCAACGACTTTCCAA	2450
QY	1240	-----	1239
DB	2451	ATAATCTACAGCGCTTTCAGCTCAGGCGTCTTAAAGCGTCTTGAAGCGCTATGGCCA	2510
QY	1240	-----	1239
DB	2511	GCTGTCTTGTGTCTTCCCTCTCACCCGCCCTGCTCTACAGCTGAGACTCCAGGAACCTT	2570
QY	1240	-----	1239
DB	2571	CAGACTACCTTCTCTGCTTCCAGAAAGGGGCTTCCCGGCTTCTCTGAGGGTCACTGG	2630
QY	1240	-----	1239
DB	2631	AGAAGCTAGACTCCCATTTGCTAGAGGTAGAAGGGGAGAGGTGTGGGAGCAGGGCTG	2690
QY	1240	-----	1239
DB	2691	GTCCACAGAGGCTCTGTCGACGAGTACCTGTGGTTCCGCCCTTCTCTATCTCTCCCTG	2750
QY	1240	-----	1239
DB	2751	TGCTCCGACCTTCCCTCCAGGCTCTGTGTATGATGCCCCCTCTCCCTCTGACGGGCTT	2810



:	APPLICANT:	Xu, Jiangchun
:	APPLICANT:	Dillon, Davin C.
:	APPLICANT:	Mitcham, Jennifer L.
:	APPLICANT:	Harlocker, Susan L.
:	APPLICANT:	Jiang, Yugu
:	APPLICANT:	Kalos, Michael D.
:	APPLICANT:	Retter, Marc W.
:	APPLICANT:	Stolk, John A.
:	APPLICANT:	Day, Craig H.
:	APPLICANT:	Vedvick, Thomas S.
:	APPLICANT:	Carter, Darrick
:	APPLICANT:	Li, Samuel X.
:	APPLICANT:	Wang, Aijun
:	APPLICANT:	Skeiky, Yasir A.W.
:	APPLICANT:	Hepler, William T.
:	APPLICANT:	Henderson, Robert A.
:	APPLICANT:	Hural, John
:	APPLICANT:	McNeill, Patricia D.
:	APPLICANT:	Houghton, Raymond L.
:	APPLICANT:	Vinals de Bassols, Carlota
:	APPLICANT:	Foy, Teresa
:	APPLICANT:	Fanger, Gary R.
:	TITLE OF INVENTION:	COMPOSITIONS AND METHODS FOR THE THERAPY AND
:	TITLE OF INVENTION:	DIAGNOSIS OF PROSTATE CANCER
:	FILE REFERENCE:	210121.534C2
:	CURRENT APPLICATION NUMBER:	US/09/895,793
:	CURRENT FILING DATE:	2001-06-29
:	NUMBER OF SEQ ID NOS:	982
:	SOFTWARE:	FastSeq for Windows Version 3.0
:	SEQ ID NO	702
:	LENGTH:	4894
:	TYPE:	DNA
:	ORGANISM:	Homo sapiens
:	US-09-895-793-702	
	Query Match	62.88; Score 2142.8; DB 9; Length 4894;
	Best Local Similarity	80.1%; Pred. No. 0;
	Matches 287;	Conservative 1; Mismatches 3; Indels 708; Gaps 2;
QY	455	GGCATTGGTCCAGTGCTGGCGCTGGTGTGTCGCCGCTCTAGGCTCAGCACGTGACCAC 514
Db	1311	GGCATTGGTCCAGTGCTGGCGCTGGTGTGTCGCCGCTCTAGGCTCAGCACGTGACCAC 1370
QY	515	TGGCGTGGACGCTATGGCCCCGGCGGCTTTCACTGTGGCAGTGTCTTGGGCATGCTG 574
Db	1371	TGGCGTGGACGCTATGGCCCCGGCGGCTTTCACTGTGGCAGTGTCTTGGGCATGCTG 1430
QY	575	CTGAGCGCTCTTTCTCATCCCAAGGGCGGCTGGCTAGCAGGGTGTGTGCCCGGATCCC 634
Db	1431	CTGAGCGCTCTTTCTCATCCCAAGGGCGGCTGGCTAGCAGGGTGTGTGCCCGGATCCC 1490
QY	635	AGSCCCCTGGAGCTGGCAGTGCTATCTCGGGCGTGGGGCTGTGTGACTTCTGTGGCCAG 694
Db	1491	AGSCCCCTGGAGCTGGCAGTGCTATCTCGGGCGTGGGGCTGTGTGACTTCTGTGGCCAG 1550
QY	695	GTGTGCTTCACTCCACTGGAGGCCCTGCTCTGTGACTCTTCCGGGACCCGGGACCACTGT 754
Db	1551	GTGTGCTTCACTCCACTGGAGGCCCTGCTCTGTGACTCTTCCGGGACCCGGGACCACTGT 1610
QY	755	CGCCAGGCTTACTTGTGTATGCTTTCAATGATCAGTCTTTGGGGCTGCCTTGGGCTACCTC 814
Db	1611	CGCCAGGCTTACTTGTGTATGCTTTCAATGATCAGTCTTTGGGGCTGCCTTGGGCTACCTC 1670
QY	815	CTGCCTGCCATTACATGGGACACAGTGCCTTGGCCCCCTACTTGGGACCCAGGAGGAG 874
Db	1671	CTGCCTGCCATTACATGGGACACAGTGCCTTGGCCCCCTACTTGGGACCCAGGAGGAG 1730
QY	875	TGCCTCTTTGGCCTGTCTCACTTCTTCTTCCCTGAGTGTGTGAGGAGGAGGAGGAG 934
Db	1731	TGCCTCTTTGGCCTGTCTCACTTCTTCTTCCCTGAGTGTGTGAGGAGGAGGAGGAG 1790
QY	935	GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 994
.	.	

QY 1369 CCCTGTGGCTGCGGTGCCACATGCTGTGCCACAGTGTGGCGGTGCACAGCTTCAGC 1428  
DB 2931 CCCTGTGGCTGCGGTGCCACATGCTGTGCCACAGTGTGGCGGTGCACAGCTTCAGC 2990  
QY 1429 CGCCCTCACGGGTTCACCTTCTACGCTTCAGCCCTGAGATCTTCCCTACACACTGCGCCCT 1488  
DB 2991 CGCCCTCACGGGTTCACCTTCTACGCTTCAGCCCTGAGATCTTCCCTACACACTGCGCCCT 3050  
QY 1489 CTACACGGGGAAGCAGGTGTCTTCCGCCAATACCGAGGGGACACTGGAGGTGCTAG 1548  
DB 3051 CTACACGGGGAAGCAGGTGTCTTCCGCCAATACCGAGGGGACACTGGAGGTGCTAG 3110  
QY 1549 CAGTGAGGACAGCTGATGACACAGCTTCTTCCGAGGCCCTTAAGCCCTGAGCTCCCTTCCC 1608  
DB 3111 CAGTGAGGACAGCTGATGACACAGCTTCTTCCGAGGCCCTTAAGCCCTGAGCTCCCTTCCC 3170  
QY 1609 TAATGAGACAGTGGGTGCTGAGGACAGTGGCTTCCACCTCCACCCGCGCTGCGG 1668  
DB 3171 TAATGAGACAGTGGGTGCTGAGGACAGTGGCTTCCACCTCCACCCGCGCTGCGG 3230  
QY 1669 GGCTCTGCTGTGATGCTCCCTAGCTGTGGTGGGTGAGCCACCGAGGCCAGGT 1728  
DB 3231 GGCTCTGCTGTGATGCTCCCTAGCTGTGGTGGGTGAGCCACCGAGGCCAGGT 3290  
QY 1729 GGTTCGGGCGGGGATCTGCTGGACCTCGCATCTCGGATGAGTGCCTTCTGCTGTC 1788  
DB 3291 GGTTCGGGCGGGGATCTGCTGGACCTCGCATCTCGGATGAGTGCCTTCTGCTGTC 3350  
QY 1789 CCAGGTGGGCCCCATCCCTGTTTATGGGCTCCATTTGCCAGCTCAGCCAGTCTGCTACTGC 1848  
DB 3351 CCAGGTGGGCCCCATCCCTGTTTATGGGCTCCATTTGCCAGCTCAGCCAGTCTGCTACTGC 3410  
QY 1849 CTATATGGTGTGCGGCGAGCCCTGGTGTGGTGGGTGAGCCACCGAGGCCAGGT 1908  
DB 3411 CTATATGGTGTGCGGCGAGCCCTGGTGTGGTGGGTGAGCCACCGAGGCCAGGT 3470  
QY 1909 ATTGACAGAGCGACTTGGCCAAATACTCAGCGTAGAAAATCTCCAGCACATTTGGGGTG 1968  
DB 3471 ATTGACAGAGCGACTTGGCCAAATACTCAGCGTAGAAAATCTCCAGCACATTTGGGGTG 3530  
QY 1969 GAGGGCTGCTCAGTGGGTCCCAGCTCCCGCTCCTGTTAGCCCCATGGGGTCCGGG 2028  
DB 3531 GAGGGCTGCTCAGTGGGTCCCAGCTCCCGCTCCTGTTAGCCCCATGGGGTCCGGG 3590  
QY 2029 CTGCGCCCGAGTTCTGCTGCCAAGTAATGTGGCTCTCTGCTGCCACCTGTGCTG 2088  
DB 3591 CTGCGCCCGAGTTCTGCTGCCAAGTAATGTGGCTCTCTGCTGCCACCTGTGCTG 3650  
QY 2089 CTGAGTGGCTGAGTGCACAGCTGGGGCTGGGGCTGCCCTCTCTCTCCCGAGTCTC 2148  
DB 3651 CTGAGTGGCTGAGTGCACAGCTGGGGCTGGGGCTGCCCTCTCTCTCCCGAGTCTC 3710  
QY 2149 TAGGGCTGCTGAGTGGAGGCTTCCAAAGGGGTTTCACTGCTGAGCTTATACAGGAGGC 2208  
DB 3711 TAGGGCTGCTGAGTGGAGGCTTCCAAAGGGGTTTCACTGCTGAGCTTATACAGGAGGC 3770  
QY 2209 CAGAAGGCTCCATGCACTGGAATGCGGGGACTCTGAGTGGATACCAGGCTCAGG 2268  
DB 3771 CAGAAGGCTCCATGCACTGGAATGCGGGGACTCTGAGTGGATACCAGGCTCAGG 3830  
QY 2269 TTAACAGTAGCTCCCTAGTTGAGACACACTAGAGAAGGTTTGGGAGCTGAATAA 2328  
DB 3831 TTAACAGTAGCTCCCTAGTTGAGACACACTAGAGAAGGTTTGGGAGCTGAATAA 3890  
QY 2329 CTCAGTCACTGGTTTCCCTCTCTAAGCCCTTAACCTCAGCTTGGTTTAATAGCT 2388  
DB 3891 CTCAGTCACTGGTTTCCCTCTCTAAGCCCTTAACCTCAGCTTGGTTTAATAGCT 3950  
QY 2389 CTTGATGGGAGTTTCTAGGATGAACACTCTCCATGGGATTTGAACATATG--ACTTA 2446  
DB 3951 CTTGATGGGAGTTTCTAGGATGAACACTCTCCATGGGATTTGAACATATGAAAGTTA 4010

QY 2447 TTTGTAGGGGAGAGTCTCTGAGGGGCAACACACAGAACACAGGTCCTCAGCCCCAGC 2506  
DB 4011 TTTGTAGGGGAGAGTCTCTGAGGGGCAACACACAGAACACAGGTCCTCAGCCCCAGC 4070  
QY 2507 ACTGTCTTTTGTGCTGATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTGTGTC 2566  
DB 4071 ACTGTCTTTTGTGCTGATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTGTGTC 4130  
QY 2567 TTTCTTGTCCATACAGAGACACAGGCAATTTAAATATTTAACTTTATTTTAAACAAGT 2626  
DB 4131 TTTCTTGTCCATACAGAGACACAGGCAATTTAAATATTTAACTTTATTTTAAACAAGT 4190  
QY 2627 AGAAGGAATCCATTGTAGCTTTTCTGTGTGTGTCTTAATATTTGGGTAGGGTGGGG 2686  
DB 4191 AGAAGGAATCCATTGTAGCTTTTCTGTGTGTGTCTTAATATTTGGGTAGGGTGGGG 4250  
QY 2687 ATCCCAACAATCAGGTCCCTGAGATGAGTGGTCTATTTGGGCTGATCATTTGCCAGAATCT 2746  
DB 4251 ATCCCAACAATCAGGTCCCTGAGATGAGTGGTCTATTTGGGCTGATCATTTGCCAGAATCT 4310  
QY 2747 TCTTCTCTGCTGGGCTCTGCGCCCAAAATGCTTAACCCAGGACCTTTGAAATTTCTACTCA 2806  
DB 4311 TCTTCTCTGCTGGGCTCTGCGCCCAAAATGCTTAACCCAGGACCTTTGAAATTTCTACTCA 4370  
QY 2807 TCCCAATATGATAATTTCCAAATGCTTTACCCAAAGTTAGGGTGTGTAAGGAAGGTAGAG 2866  
DB 4371 TCCCAATATGATAATTTCCAAATGCTTTACCCAAAGTTAGGGTGTGTAAGGAAGGTAGAG 4430  
QY 2867 GTGGGGCTTCAAGCTTCAAGCTTCCCTTAACCCAGGCTCTTCTCTTGGGCGCAGCTGGT 2926  
DB 4431 GTGGGGCTTCAAGCTTCAAGCTTCCCTTAACCCAGGCTCTTCTCTTGGGCGCAGCTGGT 4490  
QY 2927 TCCCGCCACTTCCACTCCCTCTACTCTCTAGGACTGGGCTGATGAAGGACACTGCCCA 2986  
DB 4491 TCCCGCCACTTCCACTCCCTCTACTCTCTAGGACTGGGCTGATGAAGGACACTGCCCA 4550  
QY 2987 AAATTTCCCTACCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 3046  
DB 4551 AAATTTCCCTACCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 4610  
QY 3047 TGTGTGGAGCTACTCAGGACCAAGACAGCAAAAGTGGGTTTCCCAAGCTTTGTCCATC 3106  
DB 4611 TGTGTGGAGCTACTCAGGACCAAGACAGCAAAAGTGGGTTTCCCAAGCTTTGTCCATC 4670  
QY 3107 TCAGCCCCCAGAGTATCTGTCTTGGGAACTCTCACACAGAACTCAGGAGCACCC 3166  
DB 4671 TCAGCCCCCAGAGTATCTGTCTGTTGGGAACTCTCACACAGAACTCAGGAGCACCC 4730  
QY 3167 TCCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAAAGTCCGTTTGCATAATG 3226  
DB 4731 TCCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAAAGTCCGTTTGCATAATG 4790  
QY 3227 TGTCTTATTTATTTAGGGGGTGAATTTTATCTGTAAGTGGACATCAGATATAA 3286  
DB 4791 TGTCTTATTTATTTAGGGGGTGAATTTTATCTGTAAGTGGACATCAGATATAA 4850  
QY 3287 TGTATATGTGCAAAATTTAAAGGCTTTCTTATATGTTTAAAAA 3330  
DB 4851 TGTATATGTGCAAAATTTAAAGGCTTTCTTATATGTTTAAAAA 4894

## RESULT 27

US-09-895-814-702  
; Sequence 702, Application US/09895814  
; Publication No. US20020193296A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.



```

; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 702
; LENGTH: 4894
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-895-814-702

Query Match      62.8%; Score 2142.8; DB 9; Length 4894;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2872; Conservative 1; Mismatches 3; Indels 708; Gaps 2;

QY 455 GCATTGTCCAGTGTGGGCTGGTGTGTCCTCCGCTTCTAGGCTCAGCCAGTGAACAC 514
DB 1311 GCATTGTCCAGTGTGGGCTGGTGTGTCCTCCGCTTCTAGGCTCAGCCAGTGAACAC 1370

QY 515 TGGGTGACACTATGGCCGCGCGCCCTTCATCTGGGCACTGTCTTGGGCATCTG 574
DB 1371 TGGGTGACACTATGGCCGCGCGCCCTTCATCTGGGCACTGTCTTGGGCATCTG 1430

QY 575 CTGAGCCTCTTTCTATCCCAAGGCGGCTGGCTAGCAGGGCTGTGTGCGCGGATCCC 634
DB 1431 CTGAGCCTCTTTCTATCCCAAGGCGGCTGGCTAGCAGGGCTGTGTGCGCGGATCCC 1490

QY 635 AGGCCCTTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGTGGAGCTCTGTGGCCAG 694
DB 1491 AGGCCCTTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGTGGAGCTCTGTGGCCAG 1550

QY 695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
DB 1531 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

QY 755 GCCAGGCGCTACTCTGTCTATGCTTTCATGATCAGTCTTGGGGCTGCTTGGGCTAGCTC 814
DB 1611 GCCAGGCGCTACTCTGTCTATGCTTTCATGATCAGTCTTGGGGCTGCTTGGGCTAGCTC 1670

QY 815 CTGCGTGCATTTAGCTGGGACACAGTGCCTTGGCCCGCTTACCTTGGGACCCAGAGGAG 874
DB 1671 CTGCGTGCATTTAGCTGGGACACAGTGCCTTGGCCCGCTTACCTTGGGACCCAGAGGAG 1730

QY 875 TGCCTCTTTTGGCTGCTACCCCTCATCTTCTTCACTTGCCTAGCAGCCACTGCTGGTG 934
DB 1731 TGCCTCTTTTGGCTGCTACCCCTCATCTTCTTCACTTGCCTAGCAGCCACTGCTGGTG 1790

QY 935 GCTGAGGAGGAGGCTGGGCGCCACAGGAGGAGGCTGTGGCGCCCTCTCTTG 994
DB 1791 GCTGAGGAGGAGGCTGGGCGCCACAGGAGGAGGCTGTGGCGCCCTCTCTTG 1850

QY 995 TCGCCCACTCTGCTCATGCGCGGCGGCTTGGCTTTCCGAACTTGGGCGCCCTGCTT 1054
DB 1851 TCGCCCACTCTGCTCATGCGCGGCGGCTTGGCTTTCCGAACTTGGGCGCCCTGCTT 1910

QY 1055 CCCCCTGTCACCACTGTGTGTCGCGCATGCCCGCACCCCTGCGCGGCTCTTCGTGGCT 1114

```

```

DB 1911 CCCCCTGTCACCACTGTGTGTCGCGCATGCCCGCACCCCTGCGCGGCTCTTCGTGGCT 1970
QY 1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTACAGGATTCGTGGGC 1174
DB 1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTACAGGATTCGTGGGC 2030
QY 1175 GAGGGCTGTACCAGGCGGTGCCAGAGCTGAGCGGGCACCGAGCCGAGACACTAT 1234
DB 2031 GAGGGCTGTACCAGGCGGTGCCAGAGCTGAGCGGGCACCGAGCCGAGACACTAT 2090
QY 1235 GATGA----- 1239
DB 2091 GATGAAGGTAAAGCCCTTGGCAGCCAGCAGAGGCTGTGTGGGAGCGCCACAGAGACG 2150
QY 1240 ----- 1239
DB 2151 ACACCTCGGGCTGTGTCTGGGCTGGTGGCTTCTCATCTTGGCCCGACTTCTGTGAGG 2210
QY 1240 ----- 1239
DB 2211 AAAGTGGGATGGACCCCATCTGCATACAGGGCTTCTCATGAGGTGGGAACATCTCTGCT 2270
QY 1240 ----- 1239
DB 2271 TGGGTTTTCAGGAAGGCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCTGTGCCAGTT 2330
QY 1240 ----- 1239
DB 2331 TGACAGAAGGAAGGCGGAGCTTATTCAAAGTCTAGAGGAGTGGAGGATTAAGGCTGG 2390
QY 1240 ----- 1239
DB 2391 ATTTACAGATCTGCTGTGTTCCAGCCGCACTGTGCCCTCTGCTCCCCAACAGACTTTCCAA 2450
QY 1240 ----- 1239
DB 2451 ATAATCTACCAAGCGCTTTCAGCTCAGCGCTCTAGAGCGTCTTAGAGCCTATGGCCA 2510
QY 1240 ----- 1239
DB 2511 GCTGTCTTTGTGTTTCCCTCTCAGCCCGCTGTCTCAGAGCTCAGACTCCAGGAACCTTT 2570
QY 1240 ----- 1239
DB 2571 CAGACTACCTTCTCTGCTTTCAGCAAGGGCGTTCGCCACATCTCTGAGGGTCACTGG 2630
QY 1240 ----- 1239
DB 2631 AGAACCTAGACTCCCATTTGCTAGAGGTAGAAAGGGAGAGGTGTGGGGAGCAGGGCTG 2690
QY 1240 ----- 1239
DB 2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCCTTCTCATCTCCCTGAGAC 2750
QY 1240 -----AGGCGTTGCG 1248
DB 2751 TGCTCCGACCTTCCCTCCAGGCTGTGCTCATGATGCCCTCTCTCCCTCTGACAGGCTTCG 2810
QY 1249 GATGGGACCTTGGGGCTGTTCCTGCAGTGGCCCATCTCCCTCTCTCTCTCTCTCTCTCTCT 1308
DB 2811 GATGGGACCTTGGGGCTGTTCCTGCAGTGGCCCATCTCCCTCTCTCTCTCTCTCTCTCTCT 2870
QY 1309 GGACCGCTGTGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
DB 2871 GGACCGCTGTGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930
QY 1369 CCTGTGGCTGCGGTGGCCATGCTGTCCACAGTGTGGCCGTGTGTGACAGTTCAGC 1428
DB 2931 CCTGTGGCTGCGGTGGCCATGCTGTCCACAGTGTGGCCGTGTGTGACAGTTCAGC 2990
QY 1429 CCGCCTTCACCGGTTCACTCTCTCAGGCTGAGATCTGCGCCTACACACTGGCCTCCCT 1488

```





```
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 702
; LENGTH: 4894
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-702

Query Match      62.8%; Score 2142.8; DB 10; Length 4894;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2872; Conservative 1; Mismatches 3; Indels 708; Gaps 2;

QY 455 GGCATTGGTCAGTGGGCTGTGTGTGTCCTCCGCTCTAGGCTCAGGCACTGACAC 514
    |||||
Db 1311 GGCATTGGTCAGTGGGCTGTGTGTGTCCTCCGCTCTAGGCTCAGGCACTGACAC 1370

QY 515 TGGCGTGACGCTATGGCGCGCGCGGCTTCACTGCGGCACTGTCTTGGGCACTCTG 574
    |||||
Db 1371 TGGCGTGACGCTATGGCGCGCGCGGCTTCACTGCGGCACTGTCTTGGGCACTCTG 1430

QY 575 CTGAGCCTCTTTCATCCCAAGCGCGGCTGCTAGCAGGCTGCTGCCGGATCCC 634
    |||||
Db 1431 CTGAGCCTCTTTCATCCCAAGCGCGGCTGCTAGCAGGCTGCTGCCGGATCCC 1490

QY 635 AGGCGCCCTGGAGCTGGCACTGCTCATCTGCGGCTGGGCTGTGAGCTTCTGTGGCCAG 694
    |||||
Db 1491 AGGCGCCCTGGAGCTGGCACTGCTCATCTGCGGCTGGGCTGTGAGCTTCTGTGGCCAG 1550

QY 695 GTGTGCTTCACTCCAGTGGAGGCGCTGCTCTGACCTCTTCCGGGACCGCGGACCACTGT 754
    |||||
Db 1551 GTGTGCTTCACTCCAGTGGAGGCGCTGCTCTGACCTCTTCCGGGACCGCGGACCACTGT 1610

QY 755 CGCAGGCGCTACTCTGTCTATGCTTTCATGATGCTGTGGGCTGCTGGGCTACCTC 814
    |||||
Db 1611 CGCAGGCGCTACTCTGTCTATGCTTTCATGATGCTGTGGGCTGCTGGGCTACCTC 1670

QY 815 CTGCGTGCCATTGACTGGGACACCACTGCTGCTGCCCCCTTACCTGGGACCCAGGAGGAG 874
    |||||
Db 1671 CTGCGTGCCATTGACTGGGACACCACTGCTGCTGCCCCCTTACCTGGGACCCAGGAGGAG 1730

QY 875 TGCCTCTTTGGCTTGCCTACCTCATCTTCTCAGCTGCGTAGGACCACTGCTGGTG 934
    |||||
Db 1731 TGCCTCTTTGGCTTGCCTACCTCATCTTCTCAGCTGCGTAGGACCACTGCTGGTG 1790

QY 935 GCTGAGGAGGAGCGCTGGGCGCCACCGAGCGACAGAGGCTGTGGGCTGCTGGGCTCCTTG 994
    |||||
Db 1791 GCTGAGGAGGAGCGCTGGGCGCCACCGAGCGACAGAGGCTGTGGGCTGCTGGGCTCCTTG 1850

QY 995 TCGCCCCACTGCTTCCATGCGGCGCGGCTGTGCTTTCGGGAACCTTGGGCGGCTGCTT 1054
    |||||
Db 1851 TCGCCCCACTGCTTCCATGCGGCGCGGCTGTGCTTTCGGGAACCTTGGGCGGCTGCTT 1910

QY 1055 CCGCGGTGACAGCTGTGCTGCGGATGCGCGGACCTTGGCGGCTTCTGCTGGCT 1114
    |||||
Db 1911 CCGCGGTGACAGCTGTGCTGCGGATGCGCGGACCTTGGCGGCTTCTGCTGGCT 1970

QY 1115 GAGCTGTGACGCTGGATGGCACTCATGACCTTCACTGCTGTTTACAGGATTTCTGGGGC 1174
    |||||
Db 1971 GAGCTGTGACGCTGGATGGCACTCATGACCTTCACTGCTGTTTACAGGATTTCTGGGGC 2030

QY 1175 GAGGGCTGTACAGGGCGTGCACAGCTGAGCCGGGACCGAGGCGCGGAGACACTAT 1234
    |||||
Db 2031 GAGGGCTGTACAGGGCGTGCACAGCTGAGCCGGGACCGAGGCGCGGAGACACTAT 2090

QY 1235 GATCA----- 1239
```

Db 3171 TAATGACACGCTGGGTGTTGGAGGACAGTGGCCCTGCTCCAGCCTCCACCCGCGCTCTGGGG 3230  
QY 1669 GGCTCTGCTGCTGATGCTCTCGTAGCTGTGGTGGGTGAGCCACCGAGCGCAGGGT 1728  
Db 3231 GGCTCTGCTGCTGATGCTCTCGTAGCTGTGGTGGGTGAGCCACCGAGCGCAGGGT 3290  
QY 1729 GGTTCGGGGCGGGGCGATCTGCTGACCTCGCATCTCGGATGAGTGGCTTCTCTGCTGC 1788  
Db 3291 GGTTCGGGGCGGGGCGATCTGCTGACCTCGCATCTCGGATGAGTGGCTTCTCTGCTGC 3350  
QY 1789 CAGGTGGCCCATCTGCTGTTAAGGCTCCATGTCAGCTCAGCAGTCTGCTGCTG 1848  
Db 3351 CAGGTGGCCCATCTGCTGTTAAGGCTCCATGTCAGCTCAGCAGTCTGCTGCTG 3410  
QY 1849 CTATATGCTGCTGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908  
Db 3411 CTATATGCTGCTGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3470  
QY 1909 ATTTGACAGAGCGACTTGGCCAAATPACTCAGCGTAGAATACTTCCAGCACATTTGGGTG 1968  
Db 3471 ATTTGACAGAGCGACTTGGCCAAATPACTCAGCGTAGAATACTTCCAGCACATTTGGGTG 3530  
QY 1969 GAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2028  
Db 3531 GAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3590  
QY 2029 CTGGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2088  
Db 3591 CTGGCGCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3650  
QY 2089 CTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148  
Db 3651 CTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3710  
QY 2149 TAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208  
Db 3711 TAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3770  
QY 2209 CAGAAGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2268  
Db 3771 CAGAAGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3830  
QY 2269 TTAACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328  
Db 3831 TTAACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3890  
QY 2329 CTCAGTCACTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388  
Db 3891 CTCAGTCACTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3950  
QY 2389 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2446  
Db 3951 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4010  
QY 2447 TTTGAGGGGAGAGTCTGAGGGGACACACAGAACCCAGGTCCTCAGCCACACAGC 2506  
Db 4011 TTTGAGGGGAGAGTCTGAGGGGACACACAGAACCCAGGTCCTCAGCCACACAGC 4070  
QY 2507 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2566  
Db 4071 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4130  
QY 2567 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2626  
Db 4131 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4190  
QY 2627 AGAAGGGAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2686  
Db 4191 AGAAGGGAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4250  
QY 2687 ATCCCCACAATCAGTCCCTGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2746  
Db 4251 ATCCCCACAATCAGTCCCTGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4310

QY 2747 TCTTCTCTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2806  
Db 4311 TCTTCTCTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4370  
QY 2807 TCCCAATGATAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2866  
Db 4371 TCCCAATGATAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4430  
QY 2867 GTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2926  
Db 4431 GTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4490  
QY 2927 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2986  
Db 4491 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4550  
QY 2987 AAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3046  
Db 4551 AAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4610  
QY 3047 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3106  
Db 4611 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4670  
QY 3107 TCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3166  
Db 4671 TCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4730  
QY 3167 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3226  
Db 4731 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4790  
QY 3227 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3286  
Db 4791 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4850  
QY 3287 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3330  
Db 4851 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4894

## RESULT 29

US-09-780-669-702  
; Sequence 702, Application US/09780669  
; Patent No. US2002005197A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C24  
; CURRENT APPLICATION NUMBER: US/09/780,669  
; CURRENT FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 943

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 702
; LENGTH: 4894
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-702

Query Match      62.8%; Score 2142.8; DB 10; Length 4894;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2872; Conservative 1; Mismatches 3; Indels 708; Gaps 2;

QY 455 GCGATTGGTCCAGTGGTGGGCTGGTGTGTGTCGCGCTCCCTAGGCTCAGCCAGTGCACAC 514
D 1311 GGCATTGGTCCAGTGGTGGGCTGGTGTGTGTCGCGCTCCCTAGGCTCAGCCAGTGCACAC 1370
QY 515 TGGCGTGGAGCGTATGGCGCGCGCGGCTTCATCTGGGCACTGCTCTTGGGATCTGTG 574
D 1371 TGGCGTGGAGCGTATGGCGCGCGCGGCTTCATCTGGGCACTGCTCTTGGGATCTGTG 1430
QY 575 CTGAGCCTCTTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCGCTGCTGTGCCGATCCC 634
D 1431 CTGAGCCTCTTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCGCTGCTGTGCCGATCCC 1490
QY 635 AGGCCCTCGAGCTGGCACTGCTCATCTGGCGTGGGCTGCTGGACTTCTGTGGCCAG 694
D 1491 AGGCCCTCGAGCTGGCACTGCTCATCTGGCGTGGGCTGCTGGACTTCTGTGGCCAG 1550
QY 695 GTGTGCTTCACTCCACTGAGGCGCTGCTCTGACCTCTTCCGGGACCGCGGACCACTGT 754
D 1551 GTGTGCTTCACTCCACTGAGGCGCTGCTCTGACCTCTTCCGGGACCGCGGACCACTGT 1610
QY 755 CGCCAGGCTACTCTGTCTATGCCCTTCATGATCATGCTGTGGGGCTGCTGGGCTACCTC 814
D 1611 CGCCAGGCTACTCTGTCTATGCCCTTCATGATCATGCTGTGGGGCTGCTGGGCTACCTC 1670
QY 815 CTGCTGTCATGACCTGGGACCACTGCTGCTGCGGCGCTTCTGCGGACCGCAGGAGG 874
D 1671 CTGCTGTCATGACCTGGGACCACTGCTGCTGCGGCGCTTCTGCGGACCGCAGGAGG 1730
QY 875 TGCTCTTTGGGCTGCTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 934
D 1731 TGCTCTTTGGGCTGCTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1790
QY 935 GCTGAGGAGCAGCGCTGGGCGCCACCGAGCCAGCAGAGGCGCTGTGGCGCCCTCTTG 994
D 1791 GCTGAGGAGCAGCGCTGGGCGCCACCGAGCCAGCAGAGGCGCTGTGGCGCCCTCTTG 1850
QY 995 TCGCCGCACTGCTGCTCCATGCGGCGCGCTTGGCTTTCCGCAACTGCGGCGCCCTGTT 1054
D 1851 TCGCCGCACTGCTGCTCCATGCGGCGCGCTTGGCTTTCCGCAACTGCGGCGCCCTGTT 1910
QY 1055 CCGCGGCTGCACAGCTGTGCTGCGCATGCCCGCATCCCTGCGCGGCTCTTCTGCGCT 1114
D 1911 CCGCGGCTGCACAGCTGTGCTGCGCATGCCCGCATCCCTGCGCGGCTCTTCTGCGCT 1970
QY 1115 GAGCTGTGAGCTGGATGGCACTCATGACTTTCAGCTGTTTACAGGATTTCTGCGGC 1174
D 1971 GAGCTGTGAGCTGGATGGCACTCATGACTTTCAGCTGTTTACAGGATTTCTGCGGC 2030
QY 1175 GAGGGCTGTACAGGGCTGTGCCAGAGCTGAGCGGCGCACCGAGGCGCGGAGACACTAT 1234
D 2031 GAGGGCTGTACAGGGCTGTGCCAGAGCTGAGCGGCGCACCGAGGCGCGGAGACACTAT 2090
QY 1235 GATGA----- 1239
D 2091 GATGAAGTAAAGCCTTGGCAGCAGAGGCTGTGTGGAGCGCGCCACAGAGAGCG 2150
QY 1240 ----- 1239
D 2151 ACACCTGGGCTGTGTGTGGGCTGGTGGCTCTCCATCTCTGCGGCGGACTTCTGTGAGG 2210
QY 1240 ----- 1239

2211 AAAGTGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGGAAACATCTCTGCT 2270
1240 ----- 1239
2271 TGGGTTTTCAGGAAGGCTCTGGCTGCTTAGGAGTCTGATCAGAGTCTGTTGCCAGTT 2330
1240 ----- 1239
2331 TGACAGAAGGAAGGCGGAGCTTATCAAAGTCTAGAGGAGTGGAGAGTTAAGGCTGG 2390
1240 ----- 1239
2391 ATTTTCAGATCTGCCTGTTCCAGCGGAGTGTGCCCTCTGCTCCGCCAACAGACTTTCCAA 2450
1240 ----- 1239
2451 ATAATCTCACCAGCGCTTCCAGCTCAGCGCTCTCTAGAAGCGTCTTGAAGCCTATGGCCA 2510
1240 ----- 1239
2511 GCTGTCTTGTGTCTCTCACCAGCTGCTCTCAGAGCTGAGACTTCCAGGAAACCTT 2570
1240 ----- 1239
2571 CAGACTACCTTCTCTGCCCTTTCAGCAAGGGCGTTGCCACATCTCTGAGGGTCACTGG 2630
1240 ----- 1239
2631 AAGAACTAGACTCCCATCTGCTAGAGGTAGAAAGGGAAGGCTCTGGGAGCAGGCGTG 2690
1240 ----- 1239
2691 GTCCACAGCAGGTCCTGTGAGCAGGTAACCTGTGGTTCGCGCTTCTCATCTCCCTGAGAC 2750
1240 ----- AGCGCTTCG 1248
2751 TGCTCCGACCTTCCCTCCAGGCTCTGCTGATGGCGCCCTCTCCCTCTCAGGCGCTCG 2810
1249 GATGGGAGCGCTGGGCTGTTCTGAGTGGGCACTCTCCCTGGTCTTCTCTCTGTCAT 1308
2811 GATGGGAGCGCTGGGCTGTTCTGAGTGGGCACTCTCCCTGGTCTTCTCTCTGTCAT 2870
1309 GGACCGCTGTGTGAGGATTTGGGCACTCTATTTGGCCAGTGTGGCAGCTTT 1368
2871 GGACCGCTGTGTGAGGATTTGGGCACTCTGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930
1369 CCCTGTGGCTGCCGTCACATGCTGTCCACAGTGTGGCGCTGTGAGAGCTTCAGC 1428
2931 CCCTGTGGCTGCCGTCACATGCTGTCCACAGTGTGGCGCTGTGAGAGCTTCAGC 2990
1429 CGCCCTCAGCGGCTTACCTTCTCAGCCCTGACAGTCTTCCCTTACACACTGGGCTCCCT 1488
2991 CGCCCTCAGCGGCTTACCTTCTCAGCCCTGACAGTCTTCCCTTACACACTGGGCTCCCT 3050
1489 CTACACCGGAGAGGAGTGTTCCTGCCCAATACCGAGGGGACACTGGAGTGTCTAG 1548
3051 CTACACCGGAGAGGAGTGTTCCTGCCCAATACCGAGGGGACACTGGAGTGTCTAG 3110
1549 CAGTGAGGACAGCTGTATGACAGCTTCTTCCAGGCGCTTAAAGCTTGAGCTTCCCTTCCC 1608
3111 CAGTGAGGACAGCTGTATGACAGCTTCTTCCAGGCGCTTAAAGCTTGAGCTTCCCTTCCC 3170
1609 TAATGGACAGCTGGTGTGAGGAGTGTGCTCCACCTCCACCGGCTCTCGGG 1668
3171 TAATGGACAGCTGGTGTGAGGAGTGTGCTCCACCTCCACCGGCTCTCGGG 3230
1669 GGCCTCTCCCTGTGATGCTCCGCTACGCTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1728
3231 GGCCTCTCCCTGTGATGCTCCGCTACGCTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3290
1729 GGTTCGCGGCGGGGCACTCTGCTGGAGCTCGCCATCTCGGATAGTGGCTTCTCTGCTGTC 1788
3291 GGTTCGCGGCGGGGCACTCTGCTGGAGCTCGCCATCTCGGATAGTGGCTTCTCTGCTGTC 3350
```

```
QY 1789 CCAGGTGGCCCATCCCTGTTTATGGCTCCATTGTCAGCTCAGCCAGTCTGTCACTGC 1848
DB 3351 CCAGGTGGCCCATCCCTGTTTATGGCTCCATTGTCAGCTCAGCCAGTCTGTCACTGC 3410
QY 1849 CTATATGTTGTTGCGCGCAGGCTGGGTCTGGTCGCCATTTACTTTGTACACAGGTAGT 1908
DB 3411 CTATATGTTGTTGCGCGCAGGCTGGGTCTGGTCGCCATTTACTTTGTACACAGGTAGT 3470
QY 1909 ATTTGACAGAGGACTTGGCCAAATACTCAGCGTAGAAAATCTCCAGCAATTTGGGGTG 1968
DB 3471 ATTTGACAGAGGACTTGGCCAAATACTCAGCGTAGAAAATCTCCAGCAATTTGGGGTG 3530
QY 1969 GAGGGCTCCCTCACTGGGTCCAGCTCCCGCTCCTGTTAGCCCATGGGGCTCCCGGG 2028
DB 3531 GAGGGCTCCCTCACTGGGTCCAGCTCCCTGCTCTGTTAGCCCATGGGGCTCCCGGG 3590
QY 2029 CTGGCGCGCAGTTCTGTTGCTGCGCAAGTAATGTTGGTCTGCTGCCACCCCTGTGCTG 2088
DB 3591 CTGGCGCGCAGTTCTGTTGCTGCGCAAGTAATGTTGGTCTGCTGCCACCCCTGTGCTG 3650
QY 2089 CTGAGTGTGCTAGCTGCACAGCTGGGGCTGGGGCTGCCCTCTCTCTCTCTCCCACTCTC 2148
DB 3651 CTGAGTGTGCTAGCTGCACAGCTGGGGCTGGGGCTGCCCTCTCTCTCTCTCCCACTCTC 3710
QY 2149 TAGGGCTGCTGACTGGAGGCTTCCAAGGGGTTTCACTGTGGACTTATACAGGAGGC 2208
DB 3711 TAGGGCTGCTGACTGGAGGCTTCCAAGGGGTTTCACTGTGGACTTATACAGGAGGC 3770
QY 2209 CAGAAGGGTCCATGCACTGGAAATCGGGGACTCTGCAGGTGGATACCCAGGCTCAGGG 2268
DB 3771 CAGAAGGGTCCATGCACTGGAAATCGGGGACTCTGCAGGTGGATACCCAGGCTCAGGG 3830
QY 2269 TTAACAGCTAGCTCTAGTTGAGACACACTAGAGAAGGTTTGGAGCTGTAATAA 2328
DB 3831 TTAACAGCTAGCTCTAGTTGAGACACACTAGAGAAGGTTTGGAGCTGTAATAA 3890
QY 2329 CTCACTCACTGTTTCCCATCTCTAAGCCCTTAACTGCAGCTTCGTTTAACTAGCT 2388
DB 3891 CTCACTCACTGTTTCCCATCTCTAAGCCCTTAACTGCAGCTTCGTTTAACTAGCT 3950
QY 2389 CTTGCATGGAGTTTCTAGATGAACACTCTCTCCATGGGATTTGAACATATG--ACTTA 2446
DB 3951 CTTGCATGGAGTTTCTAGATGAACACTCTCCACCTGGGATTTGAACATATGAAGTTA 4010
QY 2447 TTTCTAGGGAAGAGTCTCAGGGCAACACAGAACAGTCCCTCAGCCACACAGC 2506
DB 4011 TTTGTAGGGAAGAGTCTCAGGGCAACACAGAACAGTCCCTCAGCCACACAGC 4070
QY 2507 ACTGTCTTTTGTGATCCACCCCTCTTACCTTTTATCAGGATGGCCCTGTTGGTCC 2566
DB 4071 ACTGTCTTTTGTGATCCACCCCTCTTACCTTTTATCAGGATGGCCCTGTTGGTCC 4130
QY 2567 TTTCTGTGCAATCAGAGACAGCAGGATTTAAATTTAACTTATTTAAACAAAGT 2626
DB 4131 TTTCTGTGCAATCAGAGACAGCAGGCAITTTAAATTTAACTTATTTTAAACAAAGT 4190
QY 2627 AGAAGGAATCAATGCTAGCTTTTCTGTGTTGGTGTCTAATTTTGGGTAGGGTGGGG 2686
DB 4191 AGAAGGAATCAATGCTAGCTTTTCTGTGTTGGTGTCTAATTTTGGGTAGGGTGGGG 4250
QY 2687 ATCCCAACAAATCAGTCCCTGAGATAGCTGGTCAATTTGGGCTGATCTGTCAGAACTCT 2746
DB 4251 ATCCCAACAAATCAGTCCCTGAGATAGCTGGTCAATTTGGGCTGATCTGTCAGAACTCT 4310
QY 2747 TCTTCTCTGGGTCTGGCCCCCAAAATGCCCTAACCCAGGACCTTTGGAATTTCTACTCA 2806
DB 4311 TCTTCTCTGGGTCTGGCCCCCAAAATGCCCTAACCCAGGACCTTTGGAATTTCTACTCA 4370
QY 2807 TCCCAATGATAATTCAAATGCTGTACCAGAGTTAGGGTGTTCAGGAAGGTAGAGG 2866
DB 4371 TCCCAATGATAATTCAAATGCTGTACCAGAGTTAGGGTGTTCAGGAAGGTAGAGG 4430

RESULT 30
US-09-822-827-702
; Sequence 702, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822.827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 702
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-702

Query Match 62.8%; Score 2142.8; DB 10; Length 4894;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2872; Conservative 1; Mismatches 3; Indels 708; Gaps 2;

QY 455 GSCATTGTGCAGTGTGGGCTGTGTCTGTCGCGCTCCCTAGGCTCAGCCAGTGACCAC 514
DB 1311 GSCATTGTGCAGTGTGGGCTGTGTCTGTCGCGCTCCCTAGGCTCAGCCAGTGACCAC 1370
QY 515 TGGCGTGGACGCTATGGCCGCCGCCGCCCTTTCATCTGGGCACTGTCTCTGGGCATCTG 574
DB 1371 TGGCGTGGACGCTATGGCCGCCGCCGCCCTTTCATCTGGGCACTGTCTCTGGGCATCTG 1430
QY 575 CTGAGCCCTTCTTCTCATCCCAAGGCCCGGCTGGCTAGCAGGCTGTGTGCCCGGATCCC 634
DB 1431 CTGAGCCCTTCTTCTCATCCCAAGGCCCGGCTGGCTAGCAGGCTGTGTGCCCGGATCCC 1490
QY 635 AGGCCCTTGGAGCTGGCACTGTCTATCTGGCGTGGGGCTGTCTGGACTCTCTGTGGCCAG 694
DB 1491 AGGCCCTTGGAGCTGGCACTGTCTATCTGGCGTGGGGCTGTCTGGACTCTCTGTGGCCAG 1550
QY 695 GTGTGCTTCACTCCACTGGAGGCCCTGTCTCTGTACCTTCTTCCGGGGACCCCGGACCTGT 754
```

Db 1551 GTGTGCTTCACTCCACTGGAGCCCTGCTCTGTGACCTTTCGCGGACCCGAGCACTGT 1610  
QY 755 CGCCAGGCTACTCTGTATGCCCTTATGATCAGTCTTGGGGCTGCTGGGTACTCTC 814  
Db 1611 CGCCAGGCTACTCTGTATGCCCTTATGATCAGTCTTGGGGCTGCTGGGTACTCTC 1670  
QY 815 CTGCTGCCATTGACTGGACACCACTGCTGCCCTTACCTGGGCACCCAGAGGAG 874  
Db 1671 CTGCTGCCATTGACTGGACACCACTGCTGCCCTTACCTGGGCACCCAGAGGAG 1730  
QY 875 TGCCTCTTTGGCTGCTCACCCTCATCTCTCCTCACCCTGCGTAGACGCCACACTGCTGGTG 934  
Db 1731 TGCCTCTTTGGCTGCTCACCCTCATCTCTCCTCACCCTGCGTAGACGCCACACTGCTGGTG 1790  
QY 935 GCTGAGGAGGACGCTGGGCCCCACCGAGCCAGAGAGGCTGTCGGCCCCCTCTCTTG 994  
Db 1791 GCTGAGGAGGACGCTGGGCCCCACCGAGCCAGAGAGGCTGTCGGCCCCCTCTCTTG 1850  
QY 995 TCGCCCCACTGCTGTTCATGCGGGCCGCTGTGGCTTTCGGAACCTGGGGCCCTGCTT 1054  
Db 1851 TCGCCCCACTGCTGTTCATGCGGGCCGCTGTGGCTTTCGGAACCTGGGGCCCTGCTT 1910  
QY 1055 CCGCGCTGCACAGCTGTGCTGCGCATGCCCGCACCTCGCGCGCTTCTGTGGCT 1114  
Db 1911 CCGCGCTGCACAGCTGTGCTGCGCATGCCCGCACCTCGCGCGCTTCTGTGGCT 1970  
QY 1115 GAGCTGTGACGTGGATGGCACTCATGACCTTCACGCTTTCACGGATTTCTGTGGGC 1174  
Db 1971 GAGCTGTGACGTGGATGGCACTCATGACCTTCACGCTTTCACGGATTTCTGTGGGC 2030  
QY 1175 GAGGGCTGTACCCAGGGCTGCCAGACTGAGCGGGCCAGCGCCGAGACACTAT 1234  
Db 2031 GAGGGCTGTACCCAGGGCTGCCAGACTGAGCGGGCCAGCGCCGAGACACTAT 2090  
QY 1235 GATGA----- 1239  
Db 2091 GATGAGGTAAAGCTTGGCAGCCAGAGAGCTGGTGGGAGCCGCCACCCAGAGAGC 2150  
QY 1240 ----- 1239  
Db 2151 ACACCTCGGGGTGTGTCTGGGTGTGTGCTCTCCATCTGCGCCCGACTTCTCTCAGG 2210  
QY 1240 ----- 1239  
Db 2211 AAATGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGAACATCTGTCT 2270  
QY 1240 ----- 1239  
Db 2271 TGCCTTTTCAGGAAGGCTCTGGCTGTCTTAGAGTCTGATCAGAGTCTGTGCCCCAGTT 2330  
QY 1240 ----- 1239  
Db 2331 TGACAGAAAGGCGGAGCTTATCAAGCTTAGAGGAGTGGAGGATTAAGGCTGG 2390  
QY 1240 ----- 1239  
Db 2391 ATTTAGATCTGCTGTTCAGCGCAGTGTGCCCTCTGCTCCCCCAAGACTTTCCAA 2450  
QY 1240 ----- 1239  
Db 2451 ATAATCTCACCAGGCTTTCAGCTCAGGCTCCTAGAAGCTTGAAGCTTATGSCCA 2510  
QY 1240 ----- 1239  
Db 2511 GCTGTCTTTGTTCCCTCTCACCCTGCTGCTCAGAGCTGAGACTCCAGGAAACCTT 2570  
QY 1240 ----- 1239  
Db 2571 CAGACTACCTTCTCTGCTTACGAAGGGCGTGTGCCACATTTCTGTAGGGTCACTGG 2630  
QY 1240 ----- 1239

Db 2631 AAGAACCTAGACTCCCATCTGTAGAGGTAGAAAGGAGGCTGCTGGGAGCAGGCTG 2690  
QY 1240 ----- 1239  
Db 2691 GTCCACAGCAGGTCTCTGTCGACAGGTACCTGTGTTCGCCCTTCTCATCTCCCTGAGAC 2750  
QY 1240 ----- AGGCCCTTCG 1248  
Db 2751 TGTCCGACCTTCCCTCCAGGCTCTGTCTGATGGCCCCCTCTCCCTGTCGAGGCGTTG 2810  
QY 1249 GATGGGAGCCTGGGGCTGTCTGAGTGGCCATCTCCCTGGTCTTCTCTGTGTCTAT 1308  
Db 2811 GATGGGAGCCTGGGGCTGTCTGAGTGGCCATCTCCCTGGTCTTCTCTGTGTCTAT 2870  
QY 1309 GGACCGCTGCTGACGATTCGCGACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368  
Db 2871 GGACCGCTGCTGACGATTCGCGACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930  
QY 1369 CCCTGTGGTGGCGGTGCCACATGCCCTGTCCACAGTGTGGCGTGTGTCACAGCTTT 1428  
Db 2931 CCCTGTGGTGGCGGTGCCACATGCCCTGTCCACAGTGTGGCGTGTGTCACAGCTTT 2990  
QY 1429 CGCCCTCACCGGCTTACCTTCTCAGCCCTGTCAGATCTCCCTGTCACACTGGCCTCCCT 1488  
Db 2991 CGCCCTCACCGGCTTACCTTCTCAGCCCTGTCAGATCTCCCTGTCACACTGGCCTCCCT 3050  
QY 1489 CTACCACCGGAGAGCAGGTTCCTGCCCAATACCGAGGAGACACTGGAGGTCTAG 1548  
Db 3051 CTACCACCGGAGAGCAGGTTCCTGCCCAATACCGAGGAGACACTGGAGGTCTAG 3110  
QY 1549 CAGTGAGGACAGCCTGATGACAGCTTCTGCCAGGCCCTTAAGCCTGGAGCTCCCTTCCT 1608  
Db 3111 CAGTGAGGACAGCCTGATGACAGCTTCTGCCAGGCCCTTAAGCCTGGAGCTCCCTTCCT 3170  
QY 1609 TAATGACAGCTGGTGTGGAGCAGTGGCTGCTCCACCTCCACCGCGCTCTGCGG 1668  
Db 3171 TAATGACAGCTGGTGTGGAGCAGTGGCTGCTCCACCTCCACCGCGCTCTGCGG 3230  
QY 1669 GGCTCTGCTGTGATGTCTCCAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1728  
Db 3231 GGCTCTGCTGTGATGTCTCCAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3290  
QY 1729 GGTTCGGGCGGGGATCTGCTGACCTCGCCATCTCCCTGATGATGATGATGATGATGATG 1788  
Db 3291 GGTTCGGGCGGGGATCTGCTGACCTCGCCATCTCCCTGATGATGATGATGATGATGATG 3350  
QY 1789 CCAGTGGGCGGCTTCCCTGATGGCTCCATTTCCAGCTCAGCAGCTGTCTCACTGC 1848  
Db 3351 CCAGTGGGCGGCTTCCCTGATGGCTCCATTTCCAGCTCAGCAGCTGTCTCACTGC 3410  
QY 1849 CTATATGCTGTGCGCGAGGCTGGTCTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908  
Db 3411 CTATATGCTGTGCGCGAGGCTGGTCTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3470  
QY 1909 ATTTGACAGAGCAGCTTGGCCAAATACACAGGTAGAAACTTCCAGACACTTGGGGTG 1968  
Db 3471 ATTTGACAGAGCAGCTTGGCCAAATACACAGGTAGAAACTTCCAGACACTTGGGGTG 3530  
QY 1969 GAGGGCTGCTCAGTGGGTGCCAGCTCCCGCTCTGTTAGCCCCCATGGGCTGCGGGG 2028  
Db 3531 GAGGGCTGCTCAGTGGGTGCCAGCTCCCGCTCTGTTAGCCCCCATGGGCTGCGGGG 3590  
QY 2029 CTGGCGCCAGTTTCTGTTGCTGCCAAAGTAAATGTTGTTGCTCTGCTGCTGCTGCTGCTG 2088  
Db 3591 CTGGCGCCAGTTTCTGTTGCTGCCAAAGTAAATGTTGTTGCTCTGCTGCTGCTGCTGCTG 3650  
QY 2089 CTGAGTGGTGTAGCTGACAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 2148  
Db 3651 CTGAGTGGTGTAGCTGACAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 3710  
QY 2149 TAGGGCTGCTGACTGGAGGCTTCCAGGGGTTTCACTGTGGACTTATACAGGAGGAGC 2208  
Db 3711 TAGGGCTGCTGACTGGAGGCTTCCAGGGGTTTCACTGTGGACTTATACAGGAGGAGC 3770







Db 5412 ATAGTCCCTTCTCTGCTGCTCCAGAGTGGGCCCCATCCCTCTTTATGGGCTCCATTTGTCCAGC 5471  
QY 1830 TCAGCAGTCTCTCACTGCTATATAGGTCTCTGCGCCAGGCGTGGTCTGTGCGCAATT 1889  
Db 5472 TCAGCAGTCTCTCACTGCTATATAGGTCTCTGCGCCAGGCGTGGTCTGTGCGCAATT 5531  
QY 1890 ACTTTGCTACACAGGTAGTATTTGACAAAGAGGAGCTTGGCCAAATPACTCAGCGTAGAAAA 1949  
Db 5532 ACTTTGCTACACAGGTAGTATTTGACAAAGAGGAGCTTGGCCAAATPACTCAGCGTAGAAAA 5591  
QY 1950 CTTCCAGACATTTGGGTGGGCGCTGCCTCACTGCTGGGTCCAGCTCCCGCTCTCTGTTA 2009  
Db 5592 CTTCCAGACATTTGGGTGGGCGCTGCCTCACTGCTGGGTCCAGCTCCCGCTCTCTGTTA 5651  
QY 2010 GCCCATGGGCTGCCGGGCTGCCCGCCAGTTTCTGTTGCTGCCAAAGTATGTGGCTCT 2069  
Db 5652 GCCCATGGGCTGCCGGGCTGCCCGCCAGTTTCTGTTGCTGCCAAAGTATGTGGCTCT 5711  
QY 2070 CTGCTGCCACCTGTGCTGCTGAGTGTAGCTGCACAGCTGGGGGCTGGGGGCTGCCCT 2129  
Db 5712 CTGCTGCCACCTGTGCTGCTGAGTGTAGCTGCACAGCTGGGGGCTGGGGGCTGCCCT 5771  
QY 2130 CTCTCTCTCCCGCTCTCTAGGGCTGCTGACTGAGGCGCTTCCAGGGGCTTTCAGTC 2189  
Db 5772 CTCTCTCTCCCGCTCTCTAGGGCTGCTGACTGAGGCGCTTCCAGGGGCTTTCAGTC 5831  
QY 2190 TGGACTTATACAGGAGGCGCAGAGGGCTCCATGCACCTGGAATCGGGGACTCTGCAGGT 2249  
Db 5832 TGGACTTATACAGGAGGCGCAGAGGGCTCCATGCACCTGGAATCGGGGACTCTGCAGGT 5891  
QY 2250 GGATTAACCGAGCTCAGAGGTTAAACAGCTAGCTCTCTAGTTGAGACACACCTAGAGAAGG 2309  
Db 5892 GGATTAACCGAGCTCAGAGGTTAAACAGCTAGCTCTCTAGTTGAGACACACCTAGAGAAGG 5951  
QY 2310 TTTTGGGAGCTGAATAACTCAGTCACCTGCTGTTTCCCATCTCTAAGCCCTTAACTGTC 2369  
Db 5952 TTTTGGGAGCTGAATAACTCAGTCACCTGCTGTTTCCCATCTCTAAGCCCTTAACTGTC 6011  
QY 2370 AGCTTCGTTTAAATGCTAGCTCTGCTGATGGAGGTTTCTAGGATGAACACCTCTCCATGGA 2429  
Db 6012 AGCTTCGTTTAAATGCTAGCTCTGCTGATGGAGGTTTCTAGGATGAACACCTCTCCATGGA 6071  
QY 2430 TTTGAACATATGAC--TTATTGTTAGGGAAGAGTCTCTAGGGGCAACACACAGAAGCA 2487  
Db 6072 TTTGAACATATGAGTTATTGTTAGGGAAGAGTCTCTAGGGGCAACACACAGAAGCA 6131  
QY 2488 GGTCCCTCAGCCACACAGCTGCTTTTCTGATCCACCCCTCTTACCTTTTATCA 2547  
Db 6132 GGTCCCTCAGCCACACAGCTGCTTTTCTGATCCACCCCTCTTACCTTTTATCA 6191  
QY 2548 GGATGTCCTGTTGGTCTCTCTGCTGATCAGAGACAGGCAATTTAAATATTAA 2607  
Db 6192 GGATGTCCTGTTGGTCTCTCTGCTGATCAGAGACAGGCAATTTAAATATTAA 6251  
QY 2608 CTTATTTATTTAACAAGTAGAAGGAATCATTTGCTAGCTTTTCTGTTGGTGTCTAA 2667  
Db 6252 CTTATTTATTTAACAAGTAGAAGGAATCATTTGCTAGCTTTTCTGTTGGTGTCTAA 6311  
QY 2668 TATTTGGTAGGTGGGGATCCCCAACATCAGGTCCCTGAGATAGCTGGTCAATGGG 2727  
Db 6312 TATTTGGTAGGTGGGGATCCCCAACATCAGGTCCCTGAGATAGCTGGTCAATGGG 6371  
QY 2728 CTGATCTGCGAGAACTCTCTCTCTGCTGGGCTGCGCCCCCAAAATGCTTACCCAGG 2787  
Db 6372 CTGATCTGCGAGAACTCTCTCTCTGCTGGGCTGCGCCCCCAAAATGCTTACCCAGG 6431  
QY 2788 ACCTTGAATTTCTACTCATCCCCAAATGATAATTTCCAAATGCTGTTACCCAAAGTTAGG 2847  
Db 6432 ACCTTGAATTTCTACTCATCCCCAAATGATAATTTCCAAATGCTGTTACCCAAAGTTAGG 6491  
QY 2848 TGTTGAAGGAAGGTAGAGGTTGGGCTTCAAGCTCTCAACGGCTTCCCTTAACACCCCTCT 2907  
Db 6492 TGTTGAAGGAAGGTAGAGGTTGGGCTTCAAGCTCTCAACGGCTTCCCTTAACACCCCTCT 6551

QY 2908 TCTCTGGCCAGCCTGCTGTTCCGCCACTTCCACTCCCCCTCTACTCTCTCTAGGACTGGG 2967  
Db 6552 TCTCTGGCCAGCCTGCTGTTCCGCCACTTCCACTCCCCCTCTACTCTCTCTAGGACTGGG 6611  
QY 2968 CTGATGAAGGCACTGCGCCAAATTTCCCTTACCCCAACTTTCCCTTACCCCAACTTTC 3027  
Db 6612 CTGATGAAGGCACTGCGCCAAATTTCCCTTACCCCAACTTTCCCTTACCCCAACTTTC 6671  
QY 3028 CCACAGCTTCCACACACCTGTTTGGAGCTTACTGACAGGACCAAGACACAAAGTTCGGGT 3087  
Db 6672 CCACAGCTTCCACACACCTGTTTGGAGCTTACTGACAGGACCAAGACACAAAGTTCGGGT 6731  
QY 3088 TCCCAAGCCTTTGTCATCTCAGCCCCAGAGTATCTGCTTGGGAACTCTCACACA 3147  
Db 6732 TCCCAAGCCTTTGTCATCTCAGCCCCAGAGTATCTGCTTGGGAACTCTCACACA 6791  
QY 3148 GAAACTCAGGAGCAGCCCCCTGCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGCTTAA 3207  
Db 6792 GAAACTCAGGAGCAGCCCCCTGCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGCTTAA 6851  
QY 3208 AGTCCCTTTGCAATATGCTGCTTATTTATTTAGGGGGTGAATATTTTATCTGTA 3267  
Db 6852 AGTCCCTTTGCAATATGCTGCTTATTTATTTAGGGGGTGAATATTTTATCTGTA 6911  
QY 3268 GTGAGCAATCAGAGTATATGTTTATGCTGACAAAAATTAAGGCTTCTTATATGTTAA 3327  
Db 6912 GTGAGCAATCAGAGTATATGTTTATGCTGACAAAAATTAAGGCTTCTTATATGTTAA 6971  
QY 3328 AAAAA 3332  
Db 6972 AAAAA 6976

## RESULT 32

US-09-895-793-705  
; Sequence 705, Application US/09895793  
; Publication No. US20020192763A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuhui  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriek  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Vasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C2  
; CURRENT APPLICATION NUMBER: US/09/895,793  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 705  
; LENGTH: 6976  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-895-793-705





Db 6792 GAAACTCAGGAGCACCCCTGCTGAGCTAAGGGAGGCTTATCTCAGGGGGGTTTA 6851  
Qy 3208 AGTGGCGTTTCAATAAGTGGTCTTATTTATAGCGGCTGAATATTTTATCTAA 3267  
Db 6852 AGTGGCGTTTCAATAAGTGGTCTTATTTATAGCGGCTGAATATTTTATCTAA 6911  
Qy 3268 GTGAGCAATCAGAGTATATGTTATGTCACAAAATTAAGGCTTTCTTATATGTTAA 3327  
Db 6912 GTGAGCAATCAGAGTATATGTTATGTCACAAAATTAAGGCTTTCTTATATGTTAA 6971  
Qy 3328 AAAAA 3332  
Db 6972 AAAAA 6976

## RESULT 34

US-09-759-143-705  
; Sequence 705, Application US/09759143  
; Patent No. US200202248A1  
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Devin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C23  
; CURRENT APPLICATION NUMBER: US/09/759,143  
; CURRENT FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 934  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 705  
; LENGTH: 6976  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-759-143-705

Query Match 53.2%; Score 1815.8; DB 10; Length 6976;

Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1841; Conservative 0; Mismatches 22; Indels 2; Gaps 1;

Qy 1470 CCTACACACTGGGCTCCCTCTACCAACCGGAGAGAGAGTGTCTGCGCCAAATACCGAG 1529  
Db 5112 CTTGCTCTCTCCCTTTCTTCAACCCCTCTGCTAGGTGTCTGCGCCAAATACCGAG 5171  
Qy 1530 GGGACACTGGAGTGTAGCAGTGGAGAGAGCTGATGACAGCTTCTGCGGAGCCCTA 1589  
Db 5172 GGGACACTGGAGTGTAGCAGTGGAGAGAGCTGATGACAGCTTCTGCGGAGCCCTA 5231  
Qy 1590 AGCTGGAGTCTCCCTTCCCTAATGACAGTGGGTGCTGGAGGAGTGGCTGCTCCAC 1649  
Db 5232 AGCTGGAGTCTCCCTTCCCTAATGACAGTGGGTGCTGGAGGAGTGGCTGCTCCAC 5291  
Qy 1650 CTCACCGGCTCTCGGGGCTCTGCGGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1709  
Db 5292 CTCACCGGCTCTCGGGGCTCTGCGGCTGATGCTGATGCTGATGCTGATGCTGATGCT 5351  
Qy 1710 AGCCACCGAGGAGGAGTGGTTCGGGGGCTGCTGCGGCTGATGCTGCGGCTGCGGCT 1769  
Db 5352 AGCCACCGAGGAGGAGTGGTTCGGGGGCTGCTGCGGCTGATGCTGCGGCTGCGGCTGCG 5411

Qy 1770 ATAGTGTCTTCTGCTGCTCCAGGTGGCCCATCCCTGTTTATGGCTCCATTTGTCAGC 1829  
Db 5412 ATAGTGTCTTCTGCTGCTCCAGGTGGCCCATCCCTGTTTATGGCTCCATTTGTCAGC 5471  
Qy 1830 TCAGCCAGTGTCTGCTGCTATATGCTGCTCCGAGGCTGGGTGCTGGTGCCTATTT 1889  
Db 5472 TCAGCCAGTGTCTGCTGCTATATGCTGCTCCGAGGCTGGGTGCTGGTGCCTATTT 5531  
Qy 1890 ACTTGTCTACACAGTGTATTTTACAAAGAGGCTGGCCAAATACTCAGCCTAGAAAA 1949  
Db 5532 ACTTGTCTACACAGTGTATTTTACAAAGAGGCTGGCCAAATACTCAGCCTAGAAAA 5591  
Qy 1950 CTTCCAGCATTGGGGTGGAGGCTGCTCCTACTGCTGCTCCAGCTCCCGCTCCCTTTA 2009  
Db 5592 CTTCCAGCATTGGGGTGGAGGCTGCTCCTACTGCTGCTCCAGCTCCCGCTCCCTTTA 5651  
Qy 2010 GCGGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069  
Db 5652 GCGGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5711  
Qy 2070 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2129  
Db 5712 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5771  
Qy 2130 CTCCTCTCTCCAGTCTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2189  
Db 5772 CTCCTCTCTCCAGTCTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5831  
Qy 2190 TGGACTTATACAGGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2249  
Db 5832 TGGACTTATACAGGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5891  
Qy 2250 GGATTACCCAGGCTCAGGCTTAACAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2309  
Db 5892 GGATTACCCAGGCTCAGGCTTAACAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5951  
Qy 2310 TTTTGGAGCTGAATAAATCACTCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2369  
Db 5952 TTTTGGAGCTGAATAAATCACTCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6011  
Qy 2370 AGCTGCTTATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2429  
Db 6012 AGCTGCTTATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6071  
Qy 2430 TTTGAACATATGAC - TTTATTTGTAGGGAAGAGTCTGCTGAGGGGCAACACACAAAGACCA 2487  
Db 6072 TTTGAACATATGAAAGTTATTTGTAGGGAAGAGTCTGCTGAGGGGCAACACACAAAGACCA 6131  
Qy 2488 GGTCCCTCAGCCACACAGTGTCTTTTGTGCTGATCCACCCCTCTTACCTTTTATCA 2547  
Db 6132 GGTCCCTCAGCCACACAGTGTCTTTTGTGCTGATCCACCCCTCTTACCTTTTATCA 6191  
Qy 2548 GGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2607  
Db 6192 GGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6251  
Qy 2608 CTTATTTATTAACAAAGTGAAGGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2667  
Db 6252 CTTATTTATTAACAAAGTGAAGGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6311  
Qy 2668 TATTGGGTGGGTGGGATCCCCAACAAATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2727  
Db 6312 TATTGGGTGGGTGGGATCCCCAACAAATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6371  
Qy 2728 CTGATCATTTGCCAGATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2787  
Db 6372 CTGATCATTTGCCAGATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6431  
Qy 2788 ACCTTGGAATTTACTCATCCCAATGATATTTCCAAATGATATTTCCCAAGGTTAGGG 2847  
Db 6432 ACCTTGGAATTTACTCATCCCAATGATATTTCCCAAGGTTAGGG 6491

QY 2848 TCTTGAAGGAGGTAGAGGGTGGGGCTTCAGGTCTCAAGGGTTCCTTAACCAACCCCTCT 2907  
Db TCTTGAAGGAGGTAGAGGGTGGGGCTTCAGGTCTCAAGGGTTCCTTAACCAACCCCTCT 6551  
QY 2908 TCTCTTGGCCAGCTGGTTCCTCCCAACTTCCCTTACCTCCCTCTACTCTCTAGGACTGGG 2967  
Db TCTCTTGGCCAGCTGGTTCCTCCCAACTTCCCTTACCTCCCTCTACTCTCTAGGACTGGG 6611  
QY 2968 CTGATGAAGGCACTGCCCAAAATTTCCCTTACCTCCCAACTTTCCTTACCTCCCAACTTTC 3027  
Db CTGATGAAGGCACTGCCCAAAATTTCCCTTACCTCCCAACTTTCCTTACCTCCCAACTTTC 6671  
QY 3028 CCCACCACTCCACAACTCTTTGGAGCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3087  
Db CCCACCACTCCACAACTCTTTGGAGCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6731  
QY 3088 TCCCAAGGCTTTGTCTCATCTCAGCCCCCAGAGTATATCTGTGTTGGGGAATCTCACACA 3147  
Db TCCCAAGGCTTTGTCTCATCTCAGCCCCCAGAGTATATCTGTGTTGGGGAATCTCACACA 6791  
QY 3148 GAACTCAGGAGCACTCCCTGCTGAGCTAAGGAGGAGGAGGAGGAGGAGGAGGAGG 3207  
Db GAACTCAGGAGCACTCCCTGCTGAGCTAAGGAGGAGGAGGAGGAGGAGGAGGAGG 6851  
QY 3208 AGTGCCGTTTGCATAATGCTCTTATTTATTTAGCGGGGTGAATTTTATCTACTGTA 3267  
Db AGTGCCGTTTGCATAATGCTCTTATTTATTTAGCGGGGTGAATTTTATCTACTGTA 6911  
QY 3268 GTGACCACTCAGAGTATATTTATGTTGACAAAATTAAGGAGGAGGAGGAGGAGG 3327  
Db GTGACCACTCAGAGTATATTTATGTTGACAAAATTAAGGAGGAGGAGGAGGAGG 6971  
QY 3328 AAAAA 3332  
Db 6972 AAAAA 6976

## RESULT 35

US-09-780-669-705

; Sequence 705, Application US/09780669

; Patent No. US20020051977A1

## ; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqu  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedrick, Thomas S.  
; APPLICANT: Carter, Darick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT APPLICATION NUMBER: US/09780,669  
; CURRENT FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 943  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 705  
; LENGTH: 6976  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-780-669-705

Query Match 53.2%; Score 1815,8; DB 10; Length 6976;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1841; Conservative 0; Mismatches 22; Indels 2; Gaps 1;

QY 1470 CCTACACACTGGGCTCCCTTACCAACGGGAGAGGAGGTGTTCTTCCCTGCCCAATACCGAG 1529  
Db TCTTGAAGGAGGTAGAGGGTGGGGCTTCAGGTCTCAAGGGTTCCTTAACCAACCCCTCT 6551  
QY 1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCCTGATGACCACTTCTTCCGACGGCCCTA 1589  
Db TCTTGAAGGAGGTAGAGGGTGGGGCTTCAGGTCTCAAGGGTTCCTTAACCAACCCCTCT 6611  
QY 1590 AGCTTGAGCTCCCTTCCCTTAAATGGACAGCTGGGTGCTGGAGGAGTGGCTGCTCCAC 1649  
Db AGCTTGAGCTCCCTTCCCTTAAATGGACAGCTGGGTGCTGGAGGAGTGGCTGCTCCAC 5291  
QY 1650 CTCCACCGGGCTCTCGGGGGCTCTGCTGTGATCTCTCCSPACGTGTGGTGGTGGTG 1709  
Db CTCCACCGGGCTCTCGGGGGCTCTGCTGTGATCTCTCCSPACGTGTGGTGGTGGTG 5351  
QY 1710 AGCCCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1769  
Db AGCCCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5411  
QY 1770 ATAGTGCTTCTGCTGCTCCAGGTGGGCTCCCTGTTATATGGGCTCCATTTGTCAGC 1829  
Db ATAGTGCTTCTGCTGCTCCAGGTGGGCTCCCTGTTATATGGGCTCCATTTGTCAGC 5471  
QY 1830 TCAGGACAGTCTGCTGCTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1889  
Db TCAGGACAGTCTGCTGCTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5531  
QY 1890 ACTTTGCTACACAGTGTATTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1949  
Db ACTTTGCTACACAGTGTATTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5591  
QY 1950 CTTCCAGCACTTTGGGCTGGAGGGCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 2009  
Db CTTCCAGCACTTTGGGCTGGAGGGCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 5651  
QY 2010 GCGGCTGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCG 2069  
Db GCGGCTGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCG 5711  
QY 2070 CTGCTGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2129  
Db CTGCTGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5771  
QY 2130 CTGCTCTCTCCCGAGTCTCTAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2189  
Db CTGCTCTCTCCCGAGTCTCTAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5831  
QY 2190 TGGACTTATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2249  
Db TGGACTTATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5891  
QY 2250 GGATTAACCCAGGCTCAGGGGTTAACAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2309  
Db GGATTAACCCAGGCTCAGGGGTTAACAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCT 5951  
QY 2310 TTTTGGGAGCTGAATAAATCACTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2369  
Db TTTTGGGAGCTGAATAAATCACTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 6011  
QY 2370 AGCTTCGTTTAAATGTAGCTTTGCATGGGAGTTCCTAGGATGAAACACTCTCCATGG 2429  
Db AGCTTCGTTTAAATGTAGCTTTGCATGGGAGTTCCTAGGATGAAACACTCTCCATGG 6071  
QY 2430 TTTGAACATATGAC--TTATTTTGTAGGGAAGAGTCTGTAGGAGGAGGAGGAGGAGG 2487  
Db TTTGAACATATGAC--TTATTTTGTAGGGAAGAGTCTGTAGGAGGAGGAGGAGGAGG 6131

Qy	2488	GGTCCCCTCAGCCACAGCACTGCTCTTTTGGCTGATCCACCCCTCTTACCTTTTATCA	2547
Db	6132	GGTCCCCTCAGCCACAGCACTGCTCTTTTGGCTGATCCACCCCTCTTACCTTTTATCA	6191
Qy	2548	GGATGTGGCCCTGTTGGTCTTCTGTGTCGCATACAGACACAGGCAATTTAAATATTTAA	2607
Db	6192	GGATGTGGCCCTGTTGGTCTTCTGTGTCGCATACAGACACAGGCAATTTAAATATTTAA	6251
Qy	2608	CTTATTATTATTAAACAAAGTAGAAGGAATCCATTGCTAGCTTTTCTGTGTGGTGCCTAA	2667
Db	6252	CTTATTATTATTAAACAAAGTAGAAGGAATCCATTGCTAGCTTTTCTGTGTGGTGCCTAA	6311
Qy	2668	TATTTTGGGTAGGGTGGGGATCCCAACAATCAGGTCCCGCTGAGTAGCTGTGTCATTGGG	2727
Db	6312	TATTTTGGGTAGGGTGGGGATCCCAACAATCAGGTCCCGCTGAGTAGCTGTGTCATTGGG	6371
Qy	2728	CTGATCATTTGCCAGAACTCTTCTCTCTGGGGTCTGGCCCCCAAAATGCTTAACCCAGG	2787
Db	6372	CTGATCATTTGCCAGAACTCTTCTCTCTGGGGTCTGGCCCCCAAAATGCTTAACCCAGG	6431
Qy	2788	ACCTTGGAAATTTACTCATPCCAAAATGATPAATTCCAAATGCTTTACCCAAGTTTAGGG	2847
Db	6432	ACCTTGGAAATTTACTCATPCCAAAATGATPAATTCCAAATGCTTTACCCAAGTTTAGGG	6491
Qy	2848	TGTTGAAGGAAGCTAGAGGTTGGGCTTCAGGTCTCAACGGCTTCCCTAACACACCCCTCT	2907
Db	6492	TGTTGAAGGAAGCTAGAGGTTGGGCTTCAGGTCTCAACGGCTTCCCTAACACACCCCTCT	6551
Qy	2908	TCTCTTTGGCCACGCCTGGTTTCCCCCACTTCCACTCCCTCTACTCTCTAGGACTGGG	2967
Db	6552	TCTCTTTGGCCACGCCTGGTTTCCCCCACTTCCACTCCCTCTACTCTCTAGGACTGGG	6611
Qy	2968	CTGATGAAGCACTGCCAAAATTTCCCTTACCCCAACTTTCCCTTACCTCTCTAGGACTTTC	3027
Db	6612	CTGATGAAGCACTGCCAAAATTTCCCTTACCCCAACTTTCCCTTACCTCTCTAGGACTTTC	6671
Qy	3028	CCCACCAAGTCCACAACCTGTTTGGAGTACTGTCAGGACACAGAAGTCGCGGT	3087
Db	6672	CCCACCAAGTCCACAACCTGTTTGGAGTACTGTCAGGACACAGAAGTCGCGGT	6731
Qy	3088	TCCCAAGCCTTTGTTCATCTCAGCCCCCAGAGTATATCTGTGTTGGGAATCTCACACA	3147
Db	6732	TCCCAAGCCTTTGTTCATCTCAGCCCCCAGAGTATATCTGTGTTGGGAATCTCACACA	6791
Qy	3148	GAAACTCAGGACACCCCTGCTGAGCTAAGGGAGTCTATCTCTCAGGGGGGGTTTA	3207
Db	6792	GAAACTCAGGACACCCCTGCTGAGCTAAGGGAGTCTATCTCTCAGGGGGGGGGTTTA	6851
Qy	3208	AGTGCCGTTTGGCAATATGTCGCTTATTATTAGCGGGGTGAATATTATTACTGTAA	3267
Db	6852	AGTGCCGTTTGGCAATATGTCGCTTATTATTAGCGGGGTGAATATTATTACTGTAA	6911
Qy	3268	GTGAGCAANTCAGAGTATATGTTTATGTTGTCACAAAATTAAGGCTTCTTATATGTTTAA	3327
Db	6912	GTGAGCAANTCAGAGTATATGTTTATGTTGTCACAAAATTAAGGCTTCTTATATGTTTAA	6971
Qy	3328	AAAAA 3332	
Db	6972	AAAAA 6976	

RESULT 36

```

US-09-822-827-705
; Sequence 705, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, JiangChun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822.827
; CURRENT FILING DATE: 2001-03-28

```





QY 859 GGGACCCAGGAGGAGTGCCTTTTGGCTGCTCAACCTCATCTTCTCACCTGCGTAGC 918  
DB 882 GGGACCCAGGAGGAGTGCCTTTTGGCTGCTCAACCTCATCTTCTCACCTGCGTAGC 941  
QY 919 AGCCACACTGCTGGTGGCTGAGGAGGAGGCTGGGCGCCACCGAGCCAGCAGAGGCT 978  
DB 942 AGCCACACTGCTGGTGGCTGAGGAGGAGGCTGGGCGCCACCGAGCCAGCAGAGGCT 1001  
QY 979 GTGGGCCCCCTCTTGTGGCCCCACTGCTGTCATCGCGGGCCGCTTGGCTTTCGGAA 1038  
DB 1002 GTGGGCCCCCTCTTGTGGCCCCACTGCTGTCATCGCGGGCCGCTTGGCTTTCGGAA 1061  
QY 1039 CTGGGGGCGCTCTTCCCGGCTGCACCACTGCTGCTGCGCATCCCCCGACCCCTGGG 1098  
DB 1062 CTGGGGGCGCTCTTCCCGGCTGCACCACTGCTGCTGCGCATCCCCCGACCCCTGGG 1121  
QY 1099 CCGGCTCTGCTGGCTGAGCTGTCAGCTGGAGTGGAGTGCATGACCTTCACGCTGTTTAA 1158  
DB 1122 CCGGCTCTGCTGGCTGAGCTGTCAGCTGGAGTGGAGTGCATGACCTTCACGCTGTTTAA 1181  
QY 1159 CACGGATTTCGTGGCGGAGG 1179  
DB 1182 CACGGATTTCGTGGCGGAGTG 1202

## RESULT 38

US-09-895-793-851  
; Sequence 851, Application US/09895793  
; Publication No. US20020192763A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriack  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C2  
; CURRENT APPLICATION NUMBER: US/09/895,793  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 851  
; LENGTH: 1203  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-895-793-851

Query Match 23.3%; Score 794.6; DB 9; Length 1203;  
Best Local Similarity 99.5%; Pred. No. 4.5e-142;  
Matches 797; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 379 GCGCGAGGATCACCTATGCGCGCTCTGCTGCTGGAAGTGGGGGTAGAGAGAGTT 438  
DB 402 GCGCGAATTCACCTATGCGCGCTCTGCTGCTGGAAGTGGGGGTAGAGAGAGTT 461

QY 439 CATGACCATGTTGGTGGCATTTGGTCCAGTGTGGGCTGCTGTGTGCCGCTCCCTAGG 498  
DB 462 CATGACCATGTTGGTGGCATTTGGTCCAGTGTGGGCTGCTGTGTGCCGCTCCCTAGG 521  
QY 499 CTCAGCCAGTGAACCACTGGCGTGGACGCTATGGCCGCCGCCGCCCTTTCATCTGGGCACT 558  
DB 522 CTCAGCCAGTGAACCACTGGCGTGGACGCTATGGCCGCCGCCGCCCTTTCATCTGGGCACT 581  
QY 559 GTCCCTTGGGCATCCCTGCTGAGGCTTCTTCTATCCCAAGGGCGGCTGGCTAGCAGGCT 618  
DB 582 GTCCCTTGGGCATCCCTGCTGAGGCTTCTTCTATCCCAAGGGCGGCTGGCTAGCAGGCT 641  
QY 619 GCTGTGCCCGGATCCACAGCCCTGGAGCTGCGACTGCTCATCTGCGGCTGGGCTGCT 678  
DB 642 GCTGTGCCCGGATCCACAGCCCTGGAGCTGCGACTGCTCATCTGCGGCTGGGCTGCT 701  
QY 679 GGACTTCTGTGGCCAGGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738  
DB 702 GGACTTCTGTGGCCAGGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761  
QY 739 GGACCCGAGCACTGCTGCCAGGCTACTGCTATGCTGCTATGCTGCTATGATGATGATG 798  
DB 762 GGACCCGAGCACTGCTGCCAGGCTACTGCTATGCTGCTATGATGATGATGATGATG 821  
QY 799 CTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858  
DB 822 CTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 881  
QY 859 GGGACCCAGGAGGAGTGCCTTTTGGCTGCTCACCTCATCTTCTTCTTCTTCTTCTTCTTCT 918  
DB 882 GGGACCCAGGAGGAGTGCCTTTTGGCTGCTCACCTCATCTTCTTCTTCTTCTTCTTCTTCT 941  
QY 919 AGCCACACTGCTGGTGGCTGAGGAGGAGGCTGGGCCCCACCGAGCCAGCAGAGGCT 978  
DB 942 AGCCACACTGCTGGTGGCTGAGGAGGAGGCTGGGCCCCACCGAGCCAGCAGAGGCT 1001  
QY 979 GTGCGCCCCCTCTTGTGCGCCCCACTGCTGCTCATGCGCGGCGGCTGGCTTTCGGAA 1038  
DB 1002 GTGCGCCCCCTCTTGTGCGCCCCACTGCTGCTCATGCGCGGCGGCTGGCTTTCGGAA 1061  
QY 1039 CTGGGCGCCCTGCTTCCCGGCTGCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098  
DB 1062 CTGGGCGCCCTGCTTCCCGGCTGCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121  
QY 1099 CCGGCTCTGCTGGCTGAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCT 1158  
DB 1122 CCGGCTCTGCTGGCTGAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCT 1181  
QY 1159 CACGGATTTCGTGGCGGAGG 1179  
DB 1182 CACGGATTTCGTGGCGGAGTG 1202

## RESULT 39

US-09-895-814-851  
; Sequence 851, Application US/09895814  
; Publication No. US20020193296A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darriack  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.

```
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-851

Query Match      23.3%; Score 794.6; DB 9; Length 1203;
Best Local Similarity 99.5%; Pred. No. 4.5e-142;
Matches 797; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 379 GCGCCGAGGCATCACCTATGTGCGCCCTCTGCTGGAAGTGGGGTAGAGGAGAAGTT 438
Db      ||||| |
QY 402 GCGCAATTCATCACTATGTGCGCCCTCTGCTGGAAGTGGGGTAGAGGAGAAGTT 461
Db      ||||| |
QY 439 CATGACCATGTGCTGGGCATTGGTCCAGTGTGCTGGGCTGTGTGTCCCGTCTTAGG 498
Db      ||||| |
QY 462 CATGACCATGTGCTGGGCATTGGTCCAGTGTGCTGGGCTGTGTGTCCCGTCTTAGG 521
Db      ||||| |
QY 499 CTCAGCCAGTGACCACTGGCTGGACGTATGGCCGCGCGCCCTTTCATCTGGGCACT 558
Db      ||||| |
QY 522 CTCAGCCAGTGACCACTGGCTGGACGTATGGCCGCGCGCCCTTTCATCTGGGCACT 581
Db      ||||| |
QY 559 GTCTTGGGCATCTGCTGAGCCCTTCTTCTCATCCAGGCGCGGTGGCTAGCAGGCT 618
Db      ||||| |
QY 582 GTCTTGGGCATCTGCTGAGCCCTTCTTCTCATCCAGGCGCGGTGGCTAGCAGGCT 641
Db      ||||| |
QY 619 GCTGTGCGCGGATCCAGGCGCCCTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGCT 678
Db      ||||| |
QY 642 GCTGTGCGCGGATCCAGGCGCCCTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGCT 701
Db      ||||| |
QY 679 GGACTTCTGTGGCCAGGTGCTTCACTCCAGTGGAGCCCTGCTCTGACCTCTTCCG 738
Db      ||||| |
QY 702 GGACTTCTGTGGCCAGGTGCTTCACTCCAGTGGAGCCCTGCTCTGACCTCTTCCG 761
Db      ||||| |
QY 739 GGACCCGACCACTGCTGCCAGGCTACTGCTCTATGCTTCTATGATCATGCTTGGGG 798
Db      ||||| |
QY 762 GGACCCGACCACTGCTGCCAGGCTACTGCTCTATGCTTCTATGATCATGCTTGGGG 821
Db      ||||| |
QY 799 CTGCTGGGCTACCTCTGCTGCGCTGCACTTGAAGTGGGACACCACTGGCCCTGGCCCTACCT 858
Db      ||||| |
QY 822 CTGCTGGGCTACCTCTGCTGCGCTGCACTTGAAGTGGGACACCACTGGCCCTGGCCCTACCT 881
Db      ||||| |
QY 859 GGGCACCAGGAGGAGTGCCTTTTGGCCCTGCTCAACCTCATCTTCTCACTGCGTAGC 918
Db      ||||| |
QY 882 GGGCACCAGGAGGAGTGCCTTTTGGCCCTGCTCAACCTCATCTTCTCACTGCGTAGC 941
Db      ||||| |
QY 919 AGCCACACTGCTGTGGCTGAGGAGGAGCCCTGGGCCCCACCGACAGCAGAGGCT 978
Db      ||||| |
QY 942 AGCCACACTGCTGTGGCTGAGGAGGAGCCCTGGGCCCCACCGACAGCAGAGGCT 1001
Db      ||||| |
QY 979 GTGCGCCCTCTCTGTGCGCCCTCACTGCTGTCCATGCGCGGCGCGCTTGGCTTTCCGAA 1038
Db      ||||| |
QY 1002 GTGCGCCCTCTCTGTGCGCCCTCACTGCTGTCCATGCGCGGCGCGCTTGGCTTTCCGAA 1061
Db      ||||| |
QY 1039 CTTGGGCGCCCTCTCTGTGCGCCCTCACTGCTGTCCATGCGCGGCGCGCTTGGCTTTCCG 1098
Db      ||||| |
QY 1062 CTTGGGCGCCCTCTCTGTGCGCCCTCACTGCTGTCCATGCGCGGCGCGCTTGGCTTTCCG 1121
Db      ||||| |
QY 1099 CCGGCTCTCTGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTACGCTGTTTTA 1158
Db      ||||| |

; APPLICANT: Xu, Jlangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-851

Query Match      23.3%; Score 794.6; DB 10; Length 1203;
Best Local Similarity 99.5%; Pred. No. 4.5e-142;
Matches 797; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 379 GCGCCGAGGCATCACCTATGTGCGCCCTCTGCTGGAAGTGGGGTAGAGGAGAAGTT 438
Db      ||||| |
QY 402 GCGCAATTCATCACTATGTGCGCCCTCTGCTGGAAGTGGGGTAGAGGAGAAGTT 461
Db      ||||| |
QY 439 CATGACCATGTGCTGGGCATTGGTCCAGTGTGCTGGGCTGTGTGTCCCGTCTTAGG 498
Db      ||||| |
QY 462 CATGACCATGTGCTGGGCATTGGTCCAGTGTGCTGGGCTGTGTGTCCCGTCTTAGG 521
Db      ||||| |
QY 499 CTCAGCCAGTGACCACTGGCGTGGAGCTATGGCCCGCGCGCCCTTTCATCTGGGCACT 558
Db      ||||| |
QY 522 CTCAGCCAGTGACCACTGGCGTGGAGCTATGGCCCGCGCGCCCTTTCATCTGGGCACT 581
Db      ||||| |
QY 559 GTCTTGGGCATCTGCTGAGCCCTTCTTCTCATCCAGGCGCGGTGGCTAGCAGGCT 618
Db      ||||| |
QY 582 GTCTTGGGCATCTGCTGAGCCCTTCTTCTCATCCAGGCGCGGTGGCTAGCAGGCT 641
Db      ||||| |
QY 619 GCTGTGCGCGGATCCAGGCGCCCTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGCT 678
Db      ||||| |
QY 642 GCTGTGCGCGGATCCAGGCGCCCTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGCT 701
Db      ||||| |
QY 679 GGACTTCTGTGGCCAGGTGCTTCACTCCAGTGGAGCCCTGCTCTGACCTCTTCCG 738
Db      ||||| |
QY 702 GGACTTCTGTGGCCAGGTGCTTCACTCCAGTGGAGCCCTGCTCTGACCTCTTCCG 761
Db      ||||| |
QY 739 GGACCCGACCACTGCTGCCAGGCTACTGCTCTATGCTTCTATGATCATGCTTGGGG 798
Db      ||||| |
QY 762 GGACCCGACCACTGCTGCCAGGCTACTGCTCTATGCTTCTATGATCATGCTTGGGG 821
Db      ||||| |
QY 799 CTGCTGGGCTACCTCTGCTGCGCTGCACTTGAAGTGGGACACCACTGGCCCTGGCCCTACCT 858
Db      ||||| |
QY 822 CTGCTGGGCTACCTCTGCTGCGCTGCACTTGAAGTGGGACACCACTGGCCCTGGCCCTACCT 881
Db      ||||| |
QY 859 GGGCACCAGGAGGAGTGCCTTTTGGCCCTGCTCAACCTCATCTTCTCACTGCGTAGC 918
Db      ||||| |
QY 882 GGGCACCAGGAGGAGTGCCTTTTGGCCCTGCTCAACCTCATCTTCTCACTGCGTAGC 941
Db      ||||| |
QY 919 AGCCACACTGCTGTGGCTGAGGAGGAGCCCTGGGCCCCACCGACAGCAGAGGCT 978
Db      ||||| |
QY 942 AGCCACACTGCTGTGGCTGAGGAGGAGCCCTGGGCCCCACCGACAGCAGAGGCT 1001
Db      ||||| |
QY 979 GTGCGCCCTCTCTGTGCGCCCTCACTGCTGTCCATGCGCGGCGCGCTTGGCTTTCCGAA 1038
Db      ||||| |
QY 1002 GTGCGCCCTCTCTGTGCGCCCTCACTGCTGTCCATGCGCGGCGCGCTTGGCTTTCCGAA 1061
Db      ||||| |
QY 1039 CTTGGGCGCCCTCTCTGTGCGCCCTCACTGCTGTCCATGCGCGGCGCGCTTGGCTTTCCG 1098
Db      ||||| |
QY 1062 CTTGGGCGCCCTCTCTGTGCGCCCTCACTGCTGTCCATGCGCGGCGCGCTTGGCTTTCCG 1121
Db      ||||| |
QY 1099 CCGGCTCTCTGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTACGCTGTTTTA 1158
Db      ||||| |
```





RESULT 45  
US-09-895-793-10  
; Sequence 10, Application US/09895793  
; Publication No. US20020192763A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C2  
; CURRENT APPLICATION NUMBER: US/09/895,793  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FASTSEQ for Windows Version 3.0

```

|||||
778 TNGGNGTTCC 788

RESULT 44
US-10-012-896-10
; Sequence 10, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 9, 380, 451, 565, 582, 716, 718, 758, 762, 765, 768, 771,
; LOCATION: 779, 783
; OTHER INFORMATION: n = A,T,C or G
US-10-012-896-10

Query Match 19.7%; Score 673.4; DB 9; Length 789;
Best Local Similarity 94.2%; Pred. No. 4.8e-119;
Matches 745; Conservative 0; Mismatches 40; Indels 6; Gaps 5;

QY 1341 CAGTCATATTGGCCAGTGTGGCAGCTTCCCTGTGCTGCCGTGCCAGTGCCTGTCC 1400
Db 1 CAGTCATATTGGCCAGTGTGGCAGCTTCCCTGTGCTGCCGTGCCAGTGCCTGTCC 60

QY 1401 ACAGTGTGCCGTGGTGACAGCTTACGCCCTCACCGGTTTCACTTCTCAGCCCTGC 1460
Db 61 ACAGTGTGCCGTGGTGACAGCTTACGCCCTCACCGGTTTCACTTCTCAGCCCTGC 120

QY 1461 AGATCTGCCCTACACACTGGCCTCCCTCTACACCGGGAGAACAGAGTGTCTCTGCCCA 1520
Db 121 AGATCTGCCCTACACACTGGCCTCCCTCTACACCGGGAGAACAGAGTGTCTCTGCCCA 180

QY 1521 AATACCGAGGGGACACTGGAGTGCTAGCAGTGAGGACACCTGTATGACAGCTTCTCTC 1580
Db 181 AATACCGAGGGGACACTGGAGTGCTAGCAGTGAGGACACCTGTATGACAGCTTCTCTC 240

QY 1581 CAGSCCTTAAGCCTGGAGCTCCCTTCCCTTAATGACACGTTGGGTGCTGGAGGAGTGGCC 1640

```



```
; SEQ ID NO 10
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(789)
; OTHER INFORMATION: n = A,T,C or G
US-09-895-793-10

Query Match      19.7%; Score 673.4; DB 9; Length 789;
Best Local Similarity 94.2%; Pred. No. 4.8e-119;
Matches 745; Conservative 0; Mismatches 40; Indels 6; Gaps 5;

QY 1341 CAGTCTATTGGCCAGTGGGAGCTTTCCTGTGGTGCACATGCTGTCOC 1400
Db      |||||
QY 1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
Db      |||||
QY 1461 AGATCCTGCGCTACACACTGGCTCCCTCTACACCGGGAGAGCAGGTGTTCCTGGCCA 1520
Db      |||||
QY 1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAAGCTTCTCTGC 1580
Db      |||||
QY 1581 CAGGCCCTAAGCTGGAGCTCCCTTCCCTAATGGACAGTGGGTGCTGGAGGACAGTGCC 1640
Db      |||||
QY 1641 TGTCTCCACCTCACCCGGCTCTCGGGGGCTCTCGCTGTGATCTCTCCGTACGTGTGG 1700
Db      |||||
QY 1701 TGTGGGTGAGCCACCGAGGCCAGGGTGTTCGGGGCCGATCTGCCTGGACCTCG 1760
Db      |||||
QY 1761 CCATCTGTGATAGTCTCTCTCTGCTGCCAGGTGCCCGCCATCCCTGTTTATGGCTCCA 1820
Db      |||||
QY 1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCTATATGTTCTGTCCGCGAGCCCTGGGTCTGG 1880
Db      |||||
QY 1881 TCGCCATTACTTCTGTACACAGGTAGTATTGTGACAGAGGACCTGGCCAAATACTCAG 1940
Db      |||||
QY 1941 CGTAGAAACTTCCACGA - CATTGGGGTGGAGGCGCTGCCTCACTGGGTCCCGAGCTCCC 1998
Db      |||||
QY 1999 CGCTCCTGTAGCCCGATGGGGCTGCCGGGTGGCCCGCAGTCTTCTGTGCTGCCAAAGT 2058
Db      |||||
QY 2059 AATGTGGCTCTCTGTGCCACCTCTGCT - GCTGAGGTGCGTAGCTGCACAGCTGGGGG 2117
Db      |||||
QY 2118 TGGGGCGTCCC 2128
Db      |||||
QY 2128 TNGGNGTTC 788
```

Search completed: June 9, 2003, 22:23:03  
Job time : 609 secs



XX 09-FEB-1998; 98US-0020956.  
 PR 25-FEB-1997; 97US-0806099.  
 PR 01-AUG-1997; 97US-0904804.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Dillon DC, Xu J;  
 XX  
 DR WPI; 1998-609886/51.  
 DR N-PSDB; AAV61201.  
 XX  
 PT Polypeptides comprising immunogenic portions of prostate proteins  
 PR used in a vaccine for the treatment of prostate cancer  
 XX  
 PS Example 1; Page 82-84; 130pp; English.  
 XX  
 CC The present sequence is an immunogenic portion of a prostate tumour  
 CC protein. The immunogen, or the DNA encoding it, can be used as a  
 CC vaccine for the treatment of prostate cancer. The immunogen was  
 CC isolated from a prostate tumour cDNA library obtained by subtracting  
 CC a prostate tumour cDNA expression library with a normal tissue cDNA  
 CC library.  
 XX  
 SQ Sequence 553 AA;

Alignment Scores:  
 Pred. No.: 2,89e-208 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 19 Gaps: 0

US-09-759-143-110 (1-3410) x AAW71869 (1-553)

QY 284 ATGGTCCAGAGGCTGTGGTGTAGCCGCTGCTGGCGCACCGAAGCCAGCTCTGTGTG 343  
 DB 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
 QY 344 GTCAACCTGCTAACCTTTGGCCCTGGAGGTGTGGTGGCGGAGGATCACTATGTGGCG 403  
 DB 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTrpValPro 40  
 QY 404 CCTCTGCTGTGGAAGTGGGGGTAGAGAGAGATTTCATCACCATTGTGTGGGCTTGGT 463  
 DB 41 ProLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
 QY 464 CCAGTGTGGGCTGT 523  
 DB 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CGCTATGGCGCGCGCGCGCTTCATCTGGGCACTGTCTTGGGCACTCTGCTGAGCCTC 583  
 DB 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCAGGCGCGCTGGTGTAGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 643  
 DB 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GACCTGGCACTGCTCATCTGCGGCGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 703  
 DB 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGAGGCGCTGTCTGACCTCTTCCGGGACCGGACCACTGTGCGGAGGCC 763  
 DB 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGTCTATGCTTTCATGATCAGTCTTGGGGGCTGCTGGGCTACTCTCTGCTGCC 823  
 DB 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGGACACCAAGTGGCCCTGGCCCTTACCTTGGGACCCAGGAGGAGTGTCTTT 883

DB 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200  
 QY 884 GGCCTGCTCACCCTCATCTTCTCCTACCTGCGTAGCAGCAGCAGCTGCTGCTGAGGAG 943  
 DB 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 GCAGCGCTGGGCCCCCACCAGCAGCAGCAGAGGGCTGTGGCGCCCTCTCTCTGCGCCAC 1003  
 DB 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGCTGTCCATGCGCGGCGCGCTTGGCTTTCGGGAACCTGGCGGCTGCTTCCCGGCTG 1063  
 DB 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACCAGCTGTGTGCGCGCATGCCCGCACCTGGCGGCTCTCTGCTGGCTGAGCTGTGC 1123  
 DB 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGGCACTCATGACCTTTCACGCTGTTTACAGGATTTCTGGGCGAGGGGCTG 1183  
 DB 281 SerTrpMetAlaLeuMetThrPheThrLeuPheThrThrAspPheValGlyGluGlyLeu 300  
 QY 1184 TACCAGGCGCTGCCAGAGCTGAGCGGGCACCGAGCGCGGAGCAGACATATGATGAAGC 1243  
 DB 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
 QY 1244 GTTCGATGGGCGAGCTGGGCTGTCTCTGAGTGGCGCATCTCCCTGGCTCTCTCTCTG 1303  
 DB 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTCATGGACCGGCTGGTGTGACGATTCGGCACTCGAGCACTATTTGGCCAGTGTGGCA 1363  
 DB 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 QY 1364 GCTTTCCTGCTGGTGGCGGTCACATGCTGCTGCCAGTGGCGGCTGGGCTGGTGGTGC 1423  
 DB 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380  
 QY 1424 TCAGCGCGCTCACCGGGTTACCTTCTCAGCCCTGCAGATCTCTGCGCTACACACTGGCC 1483  
 DB 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
 QY 1484 TCCCTCTACCAACCGGAGAGCAGGTGTCTGCGCCCAATACCGAGGGGACACTGGAGGT 1543  
 DB 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
 QY 1544 GCTAGCAGTAGGACAGCCTGATCACCAGTCTTCTGCGAGCGCTTAAGCCTGAGGCTCCC 1603  
 DB 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 QY 1604 TTCCCTTAATGGACACCTGGGTGTGGAGGAGTGGCTGTCTCCACCTCCACCGCGGCTC 1663  
 DB 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeu 460  
 QY 1664 TGGCGGCGCTGCTGCTGATGCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1723  
 DB 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
 QY 1724 AGGTGTGTTCGGGCGCGGCGCATCTGCTGAGCTCCGCACTCTGAGTAGTGTCTCTG 1783  
 DB 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGTCCAGGTGGCGCGCATCCCTGTTTATGGGTCCATTTGCTCCAGCTACGCCAGTCTGTC 1843  
 DB 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTCCCTATATGTTGTCTGCGGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1903  
 DB 521 ThrAlaTyrMetValSerAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTGACAGAGCGACTTTGGCCAAATACCTACGCG 1942

Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

## RESULT 2

AAW69385

ID AAW69385 standard; Protein; 553 AA.

XX AC AAW69385;

XX 08-DEC-1998 (first entry)

XX Prostate tumour specific gene clone L1-12 protein.

XX Prostate tumour specific gene; human; prostate cancer; detection;

XX KW therapy.

XX OS Homo sapiens.

XX PN WO9837418-A2.

XX PD 27-AUG-1998.

XX PF 25-FEB-1998; 98WO-US03690.

XX PR 09-FEB-1998; 98US-0904809.

XX PR 25-FEB-1997; 97US-0806596.

XX PR 01-AUG-1997; 97US-0904809.

XX PA (CORI-) CORIXA CORP.

XX PI Dillon DC, Xu J;

XX DR WPI; 1998-480805/41.

XX DR N-PSDB; AAW58586.

XX Novel human prostate specific tumour protein and fragments - useful

PT for detecting and treating prostate cancers

XX Example 1; Page 87-89; 141pp; English.

CC This sequence is encoded by a human prostate tumour specific gene, and

CC can be used in the method of the invention. The method is for detecting

CC prostate cancer comprising contacting a biological sample with an agent

CC able to bind an immunogenic portion of a prostate protein (such as

CC this protein sequence). An antibody which binds to an immunogenic

CC portion of the prostate protein, and the method can be used to detect,

CC monitor progression of, or treat prostate cancers. The antibody may

CC also be conjugated to a therapeutic agent for use in therapy of prostate

CC cancers.

XX SQ Sequence 553 AA;

## Alignment Scores:

Pred. No.: 2,89e-208 Length: 553

Score: 2861.00 Matches: 553

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 44.58% Indels: 0

DB: 19 Gaps: 0

US-09-759-143-110 (1-3410) x AAW69385 (1-553)

QY 284 ATGTCACAGAGCTGTGGGTGAGCCGCTCTGCGGCACCGGAACCCAGCTCTTGTG 343

Db 1 MetValGlnArgLeuTyrPvalSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

QY 344 GTCACCTGCTAACCTTTGGCTGGAGTGTGTGGCCCGCCAGCATCACCTATGTGCCG 403

Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

QY 404 CCTCTGCTGTGGAGTGGGGTAGACGAGGTTTCATGACCATGTTGCTGGGCATTGGT 463.

Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

QY 464 CCAGTGTGGCGCTGTCTGTCTCCGCTCCTAGGCTCAGCCAGTACACCTGGCGTGA 523

Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTyrArgGly 80

QY 524 CGTATGGCCCGCCGCGCTTCATCTGGGCACTGTCTTGGGCAATCTGTGGAGCCTC 583

Db 81 ArgTyrGlyArgArgArgProPheIleTyrPheLeuSerLeuGlyIleLeuLeuSerLeu 100

QY 584 TTTCTCATCCCAAGGCGCGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGT 643

Db 101 PheLeuIleProArgAlaGlyTyrPheLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

QY 644 GAGTGTGCACTGTCTATCTTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 703

Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

QY 704 ACTCCACTGGAGGCGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 763

Db 141 ThrProLeuGluAlaLeuLeuSerAspPheArgAspProAspHisCysArgGlnAla 160

QY 764 TACTCTGTCTATCTTCTATGATCAGTCTTGGGCTGTCTGTGGGCTGTCTGTGGGCTGT 823

Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyCysLeuGlyTyrLeuLeuProAla 180

QY 824 ATTGACTGGGACACCACTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 883

Db 181 IleAspTyrAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200

QY 884 GGCTGTCTACCTCTATCTTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 943

Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

QY 944 GCAGCGTGGGCGCCCGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 1003

Db 221 AlaAlaLeuGlyProThrGluProAlaGlyLeuSerAlaProSerLeuSerProHis 240

QY 1004 TGCTGTCTACCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGT 1063

Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

QY 1064 CACAGCTGTCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCT 1123

Db 261 HisGlnLeuCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

QY 1124 AGCTGTGAGGCACTATGACCTTACGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 1183

Db 281 SerTyrMetAlaLeuMetThrPheThrLeuPheThrArgPheValGlyGluGlyLeu 300

QY 1184 TACCAGGCGTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCT 1243

Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

QY 1244 GTTCGGATGGGCGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 1303

Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

QY 1304 GTCATGGACCGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGT 1363

Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

QY 1364 GCTTCTTCTGTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 1423

Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380

QY 1424 TCAGCGCGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCT 1483

Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

QY 1484 TCCCTCTACCCACCGGAGAGCAGGTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGT 1543

Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

QY 1544 GCTACCACTGAGACAGCCTGTGTGGGCTGTCTGTGGGCTGTCTGTGGGCTGTCTGTGG 1603

Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProGlyAlaPro 440  
 QY 1604 TTCCTAATGACACGCTGGTGGTGGAGAGGAGTGGCTGCCACCTCCACCGGCTC 1663  
 Db 441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 460  
 QY 1664 TGGCGGGCCCTGCTGCTGATGCTCCGTACGTGTGGTGGTGGAGCCACCGAGGCC 1723  
 Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValGlyGluProThrGluAla 480  
 QY 1724 AGGTGTGTTCGGCGCGGGGACATCTGCTGGAGCTCCCATCTGGATAGTCTTCCTG 1783  
 Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGTCCAGGTGGCCCATCCCTGTTATGGGTCCATTTCCAGCTTCAGCCAGTTCGTC 1843  
 Db 501 LeuSerGlnValAlaProSerLeuPheMetClySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTCCCTATATGGTGTCTGCCGAGGCTGGGTCTGGTCCCATTTACTTTGCTACACAG 1903  
 Db 521 ThrAlaIleMetValSerAlaAlaGlyIleGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTTGACAGAGCGACTTGGCCAAATACTCAGCG 1942  
 Db 541 ValValPheAspLysSerAspLeuAlaIleTyrSerAla 553

## RESULT 3

AAB28527  
 ID AAB28527 standard; Protein; 553 AA.  
 AC AAB28527;

DT 07-FEB-2001 (first entry)  
 DE Protein encoded by human breast tumour cDNA clone p5015.

KW Human; breast tumour antigen; cytosolic; immunotherapy;  
 KW breast cancer; vaccine.

OS Homo sapiens.

PN W0200061756-A2.

PD 19-OCT-2000.

PF 10-APR-2000; 2000MO-US09688.

PR 09-APR-1999; 99US-0288950.

PR 02-JUL-1999; 99US-0346327.

XX (CORI-) CORIXA CORP.

XX Reed SG, Xu J, Dillon DC;

XX WPI; 2000-638568/61.

XX N-PSDB; AAC79473.

XX A novel isolated polypeptide comprising an immunogenic portion of a  
 XX breast cancer protein useful in the detection and treatment of breast  
 XX cancer.

XX Claim 2; Page 92-93; 95pp; English.

XX The present sequence is encoded by a cDNA sequence which was isolated  
 XX from a breast tumour cDNA library. It is provided in a specification  
 XX relating to compounds for immunotherapy and diagnosis of breast cancer.  
 XX Breast tumour antigens and the polynucleotides that encode them may be  
 XX used in the production of a pharmaceutical composition to be used in the  
 XX treatment of breast cancer. Proliferated T cells and incubated antigen  
 XX presenting cells are also required. The polypeptides and polynucleotides  
 XX may also be used to produce a vaccine.

SQ Sequence 553 AA;

## Alignment Scores:

Pred. No.: 2,89e-208 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44,58% Indels: 0  
 DB: 21 Gaps: 0

US-09-759-143-110 (1-3410) x AAB28527 (1-553)

QY 284 ATGTGTCAGAGGCTGTGGGTGAGCGGCTGTGCGGACCGGAAAGCCAGCTCTGTGCTG 343  
 Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20  
 QY 344 GTCACACCTGCTAACCTTTTGGCCTGGAGGTGTGTTTGGCGGAGGCATCACCTATGTGCGG 403  
 Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40  
 QY 404 CCTCTGCTGCTGGAAGTGGGGTAGAGAGAGTTCATGACCATGCTGTGGCATTTGGT 463  
 Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
 QY 464 CCAAGTGTGGGCTGTGTGTGTCCTGCTAGGCTCAGCCAGTACACCTGTGGCGTGA 523  
 Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CGCTATGGCGCGCGCGGCTTTCATCTGCGGACTGTCTTGGGCATCTCTGAGCGCTC 583  
 Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCCAAGGGCGGCTGTGCTAGCAGGTGTGTGCTGCGGATGCCAGGCCCTGTG 643  
 Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GAGCTGGGCACTGCTCATCTCTGGGGTGGGGTGTGTGACTTCTGTGGCCAGGTGTGCTTC 703  
 Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGGAGGCGCTCTCTGACCTCTTCCGGACCGGACCGGACCTGTGCGGACGGCC 763  
 Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGTCTATGCTTCATGATCAGTCTTGGGGGCTGCTGGGCTACCTCTGCTGCTGCC 823  
 Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGGACACACAGTGGCCCTTACCTGGGACCGGACCGGAGGAGTGCCTCTTT 883  
 Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200  
 QY 884 GGCCTGTCTACCCCTCATCTTCTCACCTGGCTAGCAGCCACACTGTGTGGTGGTGGAGG 943  
 Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 GCACGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1003  
 Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGCTGTCTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1063  
 Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACAGCTGTGCTGCCCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1123  
 Db 261 HisGlnLeuGlyCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGCACTCATGACCTTTCACGCTGTTTACAGGATTTCTGGGGCGAGGGGCTG 1183  
 Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluLeu 300  
 QY 1184 TACCAGGGCGGTGCCAGAGCTGAGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1243

```
|||||
Db   301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320
QY   1244 GTTCGGATGGCAGCCCTGGGCGTCTCCCTGCAGTGGCCATCTCCCTGGTCTCTCTCTG 1303
Db   321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340
QY   1304 GTCATGAGCCGCTGGTGCAGCGATTCCGCACCTCGACAGCTCTATTGGCCAGCTGGCA 1363
Db   341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360
QY   1364 GCTTCCCTGGTGGTGGCGTGCACATGCTGTCACAGTGTGGCCGTGGTGCAGCT 1423
Db   361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380
QY   1424 TCAGCCGCTCCACCGGCTTCACCTCTCAGCCCTGCAGATCCCTGCCTACACATGGCC 1483
Db   381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400
QY   1484 TCCCTTACCACCGGAGAGAGAGTGTCTCTGCCCAATACCGAGGGGACACTGGAGGT 1543
Db   401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420
QY   1544 GCTAGCAGTAGGACAGCCTGATGACAGCTTCTGCCAGGCCCTAAGCCTGGAGCTCC 1603
Db   421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440
QY   1604 TTCCTTAATGACACAGTGGTGGTGGAGGAGTGGCTCTCCACCTCCACCGCGCTC 1663
Db   441 PheProAsnGlyHisValGlyAlaGlyCysGlySerGlyLeuLeuProProProAlaLeu 460
QY   1664 TCGGGGCGCTCGCTGTGATGTCTCCCTACGTGTGGTGGTGGTGGAGCCACCGAGGCC 1723
Db   461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480
QY   1724 AGGTGTTCGGGCGGGGCGGATCTGCTGGACCTCCCATCTCGGATAGTGCCTCTCTG 1783
Db   481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500
QY   1784 CTGTCCAGGTGGCCCATCCCTCTTTATGGGTCCATTTGTCAGCTCAGCCAGTCTGTC 1843
Db   501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520
QY   1844 ACTGCCTATATGTGTCTGCGCAGGCGTGGTCTGCTGCGCCATTTACTTTGTACACAG 1903
Db   521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540
QY   1904 GTAGTATTTGACAGAGGAGCTTGGCCAAATACTACGG 1942
Db   541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

## RESULT 4

AAV82002 standard; Protein; 553 AA.

AC AAV82002;

13-JUN-2000 (first entry)

DE Human immunogenic prostate tumour protein sequence SEQ ID NO:113.  
XX Human; prostate cancer; diagnosis; tumour; gene therapy; detection;  
KW immunogenic; cytostatic; vaccine.

OS Homo sapiens.

XX WO200004149-A2.

PN 27-JAN-2000.

XX 14-JUL-1999; 99WO-US15838.

XX 14-JUL-1998; 98US-0115453.

```
PR 14-JUL-1998; 98US-0116134.
PR 23-SEP-1998; 98US-0159812.
PR 23-SEP-1998; 98US-0159822.
PR 15-JAN-1999; 99US-0232149.
PR 15-JAN-1999; 99US-0232880.
PR 09-APR-1999; 99US-0288946.
XX (CORI-) CORIXA CORP.
XX PA
XX PI Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;
XX WPI; 2000-171268/15.
XX DR
XX PT New polypeptide useful for treating and diagnosing prostate cancer
XX PT comprises an immunogenic portion of prostate tumor protein -
XX Claim 3; Page 138-139; 263pp; English.
XX CC
XX CC The present invention describes isolated polypeptides, comprising an
XX CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
XX CC and polynucleotides encoding them have cytostatic activity and can be
XX CC used in vaccines and in gene therapy. The polypeptides and
XX CC polynucleotides encoding them, antigen presenting cells which express
XX CC the polypeptides, antibodies against the polypeptides and vaccines
XX CC comprising them can be used for inhibiting the development of prostate
XX CC cancer in a patient. The polypeptides can be used to generate antibodies
XX CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
XX CC the polynucleotides encoding the polypeptides can be used as a probe or
XX CC to modulate the expression of the polypeptides. AA06241 to AA06691 and
XX CC AAY82000 to AAY82020 represent sequences used in the exemplification of
XX CC the present invention.
XX SQ Sequence 553 AA;
```

## Alignment Scores:

Pred. No.: 2,89e-208 Length: 553  
Score: 2861.00 Matches: 553  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 44, 58% Indels: 0  
DB: 21 Gaps: 0

US-09-759-143-110 (1-3410) x AAY82002 (1-553)

```
QY 284 ATGGTCCACAGGCTGTGGGTGAGCCCTGCTGGGACCCGAAAGCCAGCTCTTGCTG 343
Db 1 MetValGlnArgLeuTyrValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20
QY 344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTGGCCGACGATCACCTATGCGG 403
Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40
QY 404 CCTCTCTCTGCTGAAGTGGGGGTAGAGAGAGTTCATCACCATGCTGCTGGCATTGGT 463
Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60
QY 464 CCAGTGTCTGGCTGTCTGTGTCGCTCCAGCTCAGGCTAGGCTGACCATGGCGTGA 523
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisThrArgGly 80
QY 524 CGCTATGGCGCGCGCGCCCTTCATCTGGGCACTGTCTCTGGGATCCTGCTGAGCCTC 583
Db 81 ArgTyrGlyArgArgArgProPheIleTyrAlaLeuSerLeuGlyIleLeuLeuSerLeu 100
QY 584 TTTCTCATCCCAAGGCCCGGCTGGCTAGCAGGCTGCTGCTGCCGATCCAGGCCCTG 643
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120
QY 644 GAGCTGSCACTGCTCATCTCTGGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 703
Db 121 GluLeuLeuLeuLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140
QY 704 ACTCCACTGGAGGCCCTGCTCTGTACCTCTTCCGGGACCCGACCACTGTGCGCAGGCC 763
```



141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 764 TACTCTCTCATGCTTTCATGATCAGTCTTTGGGGGCTGCTGGCTACCTCTGCTGCC 823  
 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 824 ATTGACTGGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883  
 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe 200  
 884 GGCCTGCTCAGCTTCT 943  
 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 944 GCAGCGCTGGGCCCCACCGAGCAGCAGAGGCTGTGCGGCCCTCTCTGTGCGCCAC 1003  
 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 1004 TGCTGTCCATGCGCGGCGCTTGGCTTTCGGAACCTGGCGGCTGCTTCCCGGCTG 1063  
 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 1064 CACAGCTGTGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123  
 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 1124 AGCTGGATGCACCTATGACCTTCACGCTGTTTACACGATTTCTGCGGCGAGGCTG 1183  
 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
 1184 TACCAGCGCTGCGCCAGAGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1243  
 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
 1244 GTTCGGATGGGACGCTGGGCTGCTTTCCTGAGTGCCTGCTGCTGCTGCTGCTGCTG 1303  
 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 1304 GTCATGACCGCTGCTGCGACGATTCGCGACCTCGAGCAGCTCTATTGGCCAGTGGCA 1363  
 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 1364 GCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423  
 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380  
 1424 TCAGCGCGCTCACCGGCTTTCATCTCTGAGCCCTGCAGCTGCTGCTGCTGCTGCTGCT 1483  
 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
 1484 TCCTCTTACCACCGGAGAGAGTGTTCCTGCCCAATACCGAGGCGACACTGGAGT 1543  
 401 SerLeuTyrHisArgGluGlyGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
 1544 GCTAGCAGTGAGGACAGCTGATGACAGCTTCTGCGGCGCTTAAGCTGAGCTGCC 1603  
 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 1604 TTCCTTAATGACAGCTGGTGTGGAGGAGTGGCTGCTGCCACCTCCACCGGCTG 1663  
 441 PheProAsnGlyHisValGlyAlaGlyCysSerGlyLeuLeuProProProAlaLeu 460  
 1664 TGGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1723  
 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
 1724 AGGGTGTTCGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1783  
 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 1784 CTCTCCAGGTGCCCGCTCTCTGTTTATGGCTTCATGTCTCAGCTCAGCGAGTGTGTC 1843

Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTGCTATATGCTCTCTCCGAGGCTGGTCTGGTGGCTGCTGCTGCTGCTGCTGCT 1903  
 Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTGTGACAGAGGACTTGCCCAANTACTCAGCG 1942  
 Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553  
 RESULT 5  
 AAU69763  
 ID .AAU69763 standard; Protein; 553 AA.  
 XX AAU69763;  
 XX 30-JAN-2002 (first entry)  
 DE Human prostate cDNA encoded protein #3.  
 XX Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 PN WO200173032-A2.  
 XX PD 04-OCT-2001.  
 XX 27-MAR-2001; 2001WO-US09919.  
 XX 27-MAR-2000; 2000US-0536857.  
 PR 09-MAY-2000; 2000US-0568100.  
 PR 12-MAY-2000; 2000US-0570737.  
 PR 13-JUN-2000; 2000US-0593793.  
 PR 27-JUN-2000; 2000US-0605783.  
 PR 10-AUG-2000; 2000US-0636215.  
 PR 29-SEP-2000; 2000US-0651236.  
 PR 06-OCT-2000; 2000US-0657279.  
 PR 02-OCT-2000; 2000US-0679426.  
 PR 10-OCT-2000; 2000US-0685166.  
 XX (CORI-) CORIXA CORP.  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skelky YAW, Hepler WT, Henderson RA;  
 XX WPI; 2001-639232/73.  
 DR N-PSDB; AAS63557.  
 XX New human prostate-specific polypeptides and polynucleotides useful for  
 the diagnosis and treatment of cancer, especially prostate cancer -  
 PS Claim 2; Page 269-270; 579pp; English.  
 XX The invention relates to isolated prostate-specific  
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,  
 CC antibodies raised against the polypeptides (or antigenic epitopes  
 CC derived from them) and antigen-presenting cells expressing the  
 CC polypeptides. The antibodies are useful for detecting the presence of  
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and  
 CC the antigen-presenting cells are useful for stimulating and/or expanding  
 CC T cells specific for a tumour protein, and for inhibiting the development  
 CC of cancer especially prostate cancer. Compositions comprising the  
 CC polynucleotide and/or polypeptide are useful for stimulating an immune  
 CC response, and for treating cancer. The oligonucleotide is useful for  
 CC detecting cancer. The present sequence is a prostate specific  
 CC polypeptide of the invention.  
 XX SQ Sequence 553 AA;

Alignment Scores:  
 Pred. No.: 2.89e-208 Length: 553

Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAU69763 (1-553)

QY 284 ATGGTCACAGGCTGGGTGGAGCCCTGCTGGCGCACCGGAAAGCCAGCTCTTGCTG 343  
 DB 1 MetValGlnArgLeuTrpValSerArgLeuLeuAraGHisArgLysAlaGlnLeuLeuLeu 20  
 QY 344 GTCACACTGCTAACCTTTGGCTGGAGGTGTGTTGGCGCAGGACATCACCTATGTGCGC 403  
 DB 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyLeuTrpValPro 40  
 QY 404 CTTCTCTGCTGGAGTGGGTAGAGAGAGATTCATGACCATGCTGCTGGCATTTGTT 463  
 DB 41 ProLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyLeuGly 60  
 QY 464 CCAGTCTGGGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 523  
 DB 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CGCTATGGCGCGCGCGCTTCATCFCIGGGCACTGTCTCTGGGCATCTCTGTGAGCCTC 583  
 DB 81 ArgTyrGlyArgArgProPheLeuTrpAlaLeuSerLeuGlyLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 643  
 DB 101 PheLeuLeuProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 703  
 DB 121 GluLeuAlaLeuLeuLeuLeuValGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGAGGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763  
 DB 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823  
 DB 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883  
 DB 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200  
 QY 884 GGCGCTGCTACCT 943  
 DB 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 GCAGCGCTGGCGCCACCGAGCCAGAGGCTGCTGCGCCCTCTCTCTCTCTCTCTCTCT 1003  
 DB 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGCTGTCTCAATGCGCGCGCTTGGCTTTCCGGAACCTGGCGCCCTGCTGCTGCTGCTG 1063  
 DB 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACCAGCTGTGCTGCGCGACGCGCCGACCCCTGCGCGGCTCTCTCTGCTGCTGCTGCT 1123  
 DB 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGCATCATGACCTTCAGCTGTGTTTTACAGGATTTCTGGCGGAGGGCTG 1183  
 DB 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
 QY 1184 TACCAGGCGCTGCGCGAGCTGAGCGCGCGCACCGAGCGCGGAGACACTATGATGAAGCC 1243  
 DB 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTrpAspGly 320  
 QY 1244 GTTCGGATGGGCGCTGGGGCTGTTCTGCTGAGTGGGCCATCTCTCTCTCTCTCTCTG 1303

DB 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTCATGACCGCGCTGGTGGCAGGATTGGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363  
 DB 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 QY 1364 GCTTTTCCCTGCTGGCTGGCGTGGCCACATGCTGCTGCCACAGTGTGGCGTGGTACAGCT 1423  
 DB 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380  
 QY 1424 TCAGCGCGCTCACCGGTTCCACCTTCTCAGCCCTCGAGATCTGCGCTCACACTGGCC 1483  
 DB 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
 QY 1484 TCCCTCTACCAACCGGAGAGCAGGTGCTGCTGCCCAATAATACCGAGGGACACTGGAGGT 1543  
 DB 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
 QY 1544 GCTAGCAGTGGAGGACAGCTGATGACCACTGCTGCTGCCAGCCCTTAAGCCTGGAGCTCCC 1603  
 DB 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 QY 1604 TTCCTTAATGACAGCTGGGTGCTGGAGGAGTGGCTGCTGCCACCTCCACCCGCGCTC 1663  
 DB 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeu 460  
 QY 1664 TCGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1723  
 DB 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
 QY 1724 AGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1783  
 DB 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGTCCTCAGTGGCGCCCATCCCTGTTATGGGCTCCATTGTCCAGCTCAGCCACTCTGTC 1843  
 DB 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1903  
 DB 521 ThrAlaTyrMetValSerAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTGACAAGAGCAGCTGGCCAAATACTCAGCG 1942  
 DB 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 6  
 AAU04961  
 ID AAU04961 standard; Protein; 553 AA.  
 XX  
 AC AAU04961;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human prostate tumour protein L1-12.  
 XX  
 KW Human; prostate tumour protein; prostate cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6262245-B1.  
 XX  
 PD 17-JUL-2001.  
 XX  
 PF 25-FEB-1998; 98US-0030607.  
 XX  
 PR 25-FEB-1997; 97US-0806099.  
 PR 01-AUG-1997; 97US-0904804.  
 PR 09-FEB-1998; 98US-0020956.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX

PI Xu J, Dillon DC;  
 XX WPI: 2001-440862/47.  
 DR N-PSDB; AAS10108.  
 XX  
 PT Novel polynucleotide encoding polypeptide comprising a portion of  
 PT prostate tumour protein useful for inhibiting development of prostate  
 PT cancer or for treating prostate cancer in a patient  
 XX  
 PS Example 1; Column 125-127; 105pp; English.  
 XX  
 CC The sequence is a partial prostate tumour protein, encoded by a prostate  
 CC tumour specific cDNA. The DNA is useful for inhibiting the development  
 CC of prostate cancer or for treating prostate cancer in a patient.  
 XX  
 SQ Sequence 553 AA;

Alignment Scores:  
 Pred. No.: 2,89e-208 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAU04961 (1-553)

QY 284 ATGGTCCAGAGGCTGTGGGTGAGCGCGCTGCTGCGGCACCGGAAAGCCAGCTCTTGCTG 343  
 Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

QY 344 GTCACCTGCTAACTTTGGCTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 403  
 Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

QY 404 CCTGCTGCTCGAAGTGGGGTGGAGAGAGTTCATGACCATGGTGGTGGTGGTGGTGGTGGTGGT 463  
 Db 41 ProLeuLeuLeuGluValGlyValGluGlyLysPheMetThrMetValLeuGlyIleGly 60

QY 464 CAGTGTGGGCTGT 523  
 Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTyrArgGly 80

QY 524 CGCTATGCGCGCGCGCGCTTCTATCTGGGCACCTGCTTGGGCATCTGCTGCTGCTGCTGCTGCTG 583  
 Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuLeuSerLeu 100

QY 584 TTTCTCATCCCAAGCGCGCTGGCTACAGGGTGTGTGCGCGGATCCAGGCGCGCTG 643  
 Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

QY 644 GAGCTGGGACCTGCTATCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 703  
 Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

QY 704 ACTCCACTGGAGGCGCTGCT 763  
 Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

QY 764 TACTGTGCTATGCTTTCATGATCAGTCTTGGGGCTGGCTGGGCTACCTCTGCTGCTGCTGCTG 823  
 Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

QY 824 ATTGACTGGGACACAGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 883  
 Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200

QY 884 GGCCTGCTCACCTTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 943  
 Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

QY 944 GCAGCGCTGGGCG 1003  
 Db 944 GCAGCGCTGGGCG

RESULT 7

AAU01117

ID AAU01117 standard; Protein: 553 AA.

XX

AC AAU01117;

XX

DT 04-OCT-2001 (first entry)

Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

QY 1004 TGCTGTCCATGCGCGCGCGCTTGGCTTTCGGAACCTGGCGCGCTGCTTCCCGGGCTG 1063

Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

QY 1064 CACCAGCTGTGTGCGCGCATGCGCCACCCCTGCGCGCGCTCTTCTGCTGAGCTGTGC 1123

Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaLeuLeuCys 280

QY 1124 AGCTGATGGCACTCATGACCTTCAGCTGCTGTTTACACGAGATTTCGTGGCGCGGGCTG 1183

Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

QY 1184 TACACGGGCTGCCAGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1243

Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

QY 1244 GTTCGGATGGCAGCGCTGGGCTGTCTGCTGAGTGGCGCATCTCCCTGGTCTTCTCTG 1303

Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

QY 1304 GTCATGACCGCGCTGTGCGCGCATTCGGCAGCTTCGAGCAGTCTATTGTCAGTGTGGCA 1363

Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

QY 1364 GCTTCCCTGTGCTGCTGCGGTGCCATGCCATGCCCTGCCACAGTGTGGCGCTGTGACAGCT 1423

Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380

QY 1424 TCAGCGCGCTTCACCGGTTTCACTTCTCAGCGCTGCGAGATCCTGCCCTACACACTGCC 1483

Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

QY 1484 TCCTCTACACCGCGGAGAGAGTGTTCCTGCCCAATACCGGGGACACTGGAGT 1543

Db 401 SerLeuTyrHisArgGluGlyGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

QY 1544 GCTAGCAGTGAAGACAGCTGTATGACAGCTTCTGCGCGCGCGCTTCCAGCGCGCTGAGCTGCC 1603

Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

QY 1604 TTCCTTAATGGACACGTGGTGTGCGAGCGCTGCTGCCACCTGCCACCTGCCACCGCGCTC 1663

Db 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProProAlaLeu 460

QY 1664 TCGCGGCGCTCTGCTGT 1723

Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

QY 1724 AGGTGTGCTTCCGCGCGCGCGCGCATCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1783

Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

QY 1784 CTGTCCCGCTGGCGCGCGCGCTGTTATGGCTCCATGCTCCAGCTCAGCGCTGCTGCTC 1843

Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

QY 1844 ACTGCCATATGCTGTCTGCTGCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1903

Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

QY 1904 GTAGTATTGACAGCGACTTGGCCAAATACTCAGCG 1942

Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

XX DE Human prostate-specific amino acid sequence L1-12.  
 XX KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
 KW cystostatic; gene therapy; metastasis.  
 XX OS Homo sapiens.  
 XX PN WO200151633-A2.  
 XX PD 19-JUL-2001.  
 XX PF 16-JAN-2001; 2001WO-US01574.  
 XX PR 14-JAN-2000; 2000US-0483672.  
 XX PA (CORI-) CORIXA CORP.  
 XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;  
 PI Wang A, Meagher MJ;  
 XX DR WPI; 2001-425873/45.  
 XX PT New polynucleotide encoding a prostate-specific protein, for  
 PT diagnosing, monitoring and treating prostate cancer in a patient and  
 PT for use in vaccines -  
 XX PS Claim 2; Page 267-268; 543pp; English.  
 XX CC The present invention describes polynucleotide sequences (I) which encode  
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
 CC and can be used in vaccine production and gene therapy. (I), (II),  
 CC antibodies to (II), fusion proteins comprising (II), and isolated  
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.  
 CC (I) and the antibodies are also used in the detection of cancer in a  
 CC patient. The cancer that is diagnosed or treated is particularly  
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
 CC (I) can be used for monitoring the progression of cancer in a patient.  
 CC (I) and (II) can also be used to improve diagnostic and therapeutic  
 CC methods for prostate cancer. They can indicate the level of metastasis  
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to  
 CC AAH01318 represent polynucleotide and amino acid sequences used in the  
 CC exemplification of the present invention.  
 XX SQ Sequence 553 AA;

Alignment Scores:  
 Pred. No.: 2,89e-208 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAH01117 (1-553)

QY 284 ATGGTCAGAGGCTGTGGGTAGCGCGCTGCTGGCGACCGAAGCCAGCTCTTGGCTG 343  
 Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20  
 QY 344 GTCACCTGTACCTTTGGCCCTGGAGGTGTGTTGGCGCGAGGCATACCTATGTCGG 403  
 Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTrpValPro 40  
 QY 404 CCTCTGCTGCTGGAAGTGGGGTAGAGGAAGTTCATGACCATGCTGCTGGCATTTGCT 463  
 Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
 QY 464 CCAGTGTGGGCGCTGTCTGTGTCGGCTCTAGGCTCAGCCAGTACACCTGGCGTGA 523  
 Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

QY 524 CGCTATGGCCGCCGCCCTTCTATCTGGGCACTGTCTTGGGCACTCTCTGAGCCTC 583  
 Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCCAAGGGCGGCTGCTACAGGGTGTGTGCTGCCGGATGCCAGGCCCTG 643  
 Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GAGCTGGCACTGCTCATCTGGCGCTGGGCTGTGGACTTCTGTGGCCAGGTGTGCTTC 703  
 Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGGAGGCCCTGCTCTCTGACCTTCTCGGGACCGCGGACCACTGTGCGAGGCC 763  
 Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGCTATGCTTCTCATGATCAGTCTTGGGGCTGCTGGGCTACTCTCTGCTGCC 823  
 Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGACACCACTGCTGCTGGCCCTTACCTGGGCACTGGGAGGAGTGTCTCTTT 883  
 Db 181 IleAspTrpAspThrSerAlaLeuAlaProIleLeuGlyThrGlnGluGluCysLeuLeu 200  
 QY 884 GGCCTGCTCACCTCATCTTCTCACCTGCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAG 943  
 Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 CGAGCGCTGGGCCCCACCGAGCCAGCAGAGGGCTGTGGCGCCCTCTTGTGGCCGCAC 1003  
 Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGCTGTCCATGCGCGGCCCTGCTTTCGGAACCTGGCGGCCCTGCTTCCCGGGCTG 1063  
 Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACCACTGTGCTGCGCATGCCCGACCTGCGCGCTGCTGCTGGTGGCTGAGTGTGC 1123  
 Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGGCACTCATGACCTTTCACGCTGTTTACACGATTTCTGGTGGGAGGGGTG 1183  
 Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheThrArgPheValGlyGluGlyLeu 300  
 QY 1184 TACCAGGCGCTGCCAGAGCTGAGCGGGCCACCGAGCGCCGAGACACTATGATGAAGGC 1243  
 Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTrpAspGluGly 320  
 QY 1244 GTTCGATGGGAGCGCTGGGGCTGTTCTCTGAGTGGCGCATCTCCCTGGTCTTCTCTCTG 1303  
 Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTCATGGACCGCTGCTGACGATTCGCGCACTTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363  
 Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValIleLeuAlaSerValAla 360  
 QY 1364 GCTTTCCTGTGGCTGCGCGTGCACATGCTGCTGCCACAGTGTGGCGTGGTGACAGCT 1423  
 Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValIleThrAla 380  
 QY 1424 TCAGCCGCGCTCACCGGGTTCACCTTCTCAGCCCTGAGATCTGCTGCCCTACACTGGCC 1483  
 Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
 QY 1484 TCCCTCTACCGCGGAGCAGGTGTTCTGCCCAATACCGAGGGACACTGGAGGT 1543  
 Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProTyrTrpArgGlyAspThrGlyGly 420  
 QY 1544 GCTAGCAGTGGAGCAGCCTGATCACCACTGCTGCTGCCAGGCCCTTAACCCCTGGAGCTCC 1603  
 Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 QY 1604 TTCCTANTGGACACGTGGGTGCTGGAGGCACTGGCCCTGCTCCCACTCCACCGCGCTC 1663

Db 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProAlaLeu 460  
 QY 1664 TCGGGGCTCTGCTGTGATGCTCCCTAGCTGTGGTGGGTGAGCCACCGAGGCC 1723  
 Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
 QY 1724 AGGGTGGTTCGGGGCGGCGGATCTGCTGACCTCGCCATCCCGATAGTCCCTCTG 1783  
 Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGCTCCAGGTGGCCCATCCCTGTTATGGCTCCATGTCACGCTCAGCCAGTCTGTC 1843  
 Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTGCTATATGTTCTGCTCGCGCAGGCTGGGCTCTGGTGGCCATTTACTTTGCTACACAG 1903  
 Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTGACAAAGAGGACTTGCCCAAACTACAGCG 1942  
 Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

## RESULT 8

AAG99002

ID AAG99002 standard; Protein; 553 AA.

XX AAG99002;

AC AAG99002;

DT 25-SEP-2001 (first entry)

DE Human prostate-specific amino acid sequence L1-12/P501S.

XX Human; prostate cancer; therapy; diagnosis; cat eye syndrome;

KW chromosome 22q11.2; prostate-specific protein; chromosome 1;

KW prostate specific antigen; PSA.

XX Homo sapiens.

OS WO200134802-A2.

PN 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30904.

XX 12-NOV-1999; 99US-0439313.

PR 18-NOV-1999; 99US-0443686.

XX (CORI-) CORIXA CORP.

PA Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;

PI PI Kalos MD, Retter MW, Stoik JA, Day CH, Skeiky YAW, Wang A;

XX WPI; 2001-308785/32.

DR Isolated polypeptide comprising at least an immunogenic portion of a

PT prostate-specific protein, useful in the diagnosis and therapy of

PT prostate cancer -

XX Claim 3; Page 167-168; 325pp; English.

PS The present invention describes an isolated polypeptide (P1) comprising

XX at least an immunogenic portion of a prostate-specific protein, or its

CC variant. Also described are polynucleotides (N1) encoding (P1). (P1) and

CC (N1) have cytostatic activity and can be used in vaccine production.

CC The polypeptides, nucleic acids and antibodies from the present

CC invention are useful in the diagnosis and therapy of prostate cancer.

CC Prostate specific genes p704P, p712P, p774P, p775P and p305P are located

CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome

CC region. Prostate specific antigen (PSA) P501S was located on

CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent

CC polynucleotide and polypeptide sequences used in the exemplification

CC of the present invention.

XX SQ Sequence 553 AA;  
 Alignment Scores:  
 Pred. No.: 2,89e-208 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAG99002 (1-553)

QY 284 ATGTCACAGAGGCTGTGGGTGAGCCGCTCTCGGCACCGGAAAGCCAGCTCTTGTCTG 343  
 Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20  
 QY 344 GTCAACCTGCTAACCTTTGGCCCTGGAGGTGTGTTGGCCGACGAGCATACCATATGTGCG 403  
 Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40  
 QY 404 CCTCTGCTGTGGAAGTGGGGGTAGAGGAGAGTTCATGACCATGCTGTGGCATTTGGT 463  
 Db 41 ProLeuLeuLeuGluValGlyValGlyValGlyLysPheMetThrMetValLeuGlyIleGly 60  
 QY 464 CCAGTCTGGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 523  
 Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CGCTATGGCGCGCGCGGCTTTCATCTGGGCACTGCTGCTGGGCACTGCTGCTGGGCACT 583  
 Db 81 ArgTyrGlyArgArgArgPropheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTATCCAAAGGCGGCTGGCTAGCAGGCTGCTGTGCTGGGCACTGCTGCTGGGCACT 643  
 Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuLeuLeuLeuLeuLeuLeu 120  
 QY 644 GAGCTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703  
 Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 763  
 Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGCTATCCCTTCATGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823  
 Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGACACACAGTGGCCCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 883  
 Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe 200  
 QY 884 GGCCTGCTCACCTTCATCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943  
 Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 GCAGCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1003  
 Db 221 AlaAlaLeuGlyProThrGluProAlaGluLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGGTGTCCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1063  
 Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1123  
 Db 261 HisGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 280  
 QY 1124 AGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1183  
 Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlyLeu 300



181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlnGluGluCysLeuPhe 200  
Db  
884 GGCCTGCTCACCTCATCTTCCTACCTCGCTAGCAGCCACACTGCTGCTGCTCAGGAG 943  
QY  
201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGlu 220  
Db  
944 GCAGCGCTGGGCCACACCGAGCCAGCAGAGGCTGCTGGCCCTCTCTTGTGCGCCAC 1003  
QY  
221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
Db  
1004 TGTGTCCATGGCGGCGCGCTTGGCTTTCGGAACCTGGGCGCCCTGCTTCCCGCGGTG 1063  
QY  
241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
Db  
1064 CACACGCTGTGTCGCGCATGCCCGCACCTGCGCGCGCTCTCTGTCGTCAGCTGTGC 1123  
QY  
261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
Db  
1124 AGCTGGATGGCACTCATGACCTTACGCTGTGTTTACACGGATTTCGTGGCGAGGGGTG 1183  
QY  
281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
Db  
1184 TACAGGGGTGCCAGAGCTGAGCGCGCACCGAGCGCCGAGACACTATGATGAAGC 1243  
QY  
301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgHisTyrAspGluGly 320  
Db  
1244 GTTCGGATGGCAGCGCTGGGCTGTTCCTGCGAGTCGCCATCTCCCTGCTCTCTCTG 1303  
QY  
321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
Db  
1304 GTCATGACCGGTGTGTCAGCGATTTCGGACATCGAGCAGTCTATTGGCCAGTGTGGA 1363  
QY  
341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
Db  
1364 GCTTCCCTGCTGCTGCGGTGCGACATGCTGCTCCACAGTGTGCGCTGTGACAGCT 1423  
QY  
361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380  
Db  
1424 TCAGCCCGCTCACCGGTTTACCTTCTCAGCCCTGCATCTGCGCTACACATCGGCG 1483  
QY  
381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
Db  
1484 TCCCTTACACCGGGAGACAGGTGTCCTGCCCAATACGAGGGACACTGAGCT 1543  
QY  
401 SerLeuTyrHisArgGluGlyGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
Db  
1544 GCTAGCAGTGGAGACGCTGATCACAGCTTCTGCGAGCGCTTAAAGCTGAGCTCCC 1603  
QY  
421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
Db  
1604 TTCCTTAATGACACGTGGGTGCTGGAGGAGTGGCTGCTCCACCTCCACCGCGCTC 1663  
QY  
441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeu 460  
Db  
1664 TGGCGGCGCTGCTGCTGATGCTCTCTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1723  
QY  
461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
Db  
1724 AGGTGCTTCCGGCGCGGCGCATCTGCTGGACCTCGCCATCTCGGATAGTGCCTCTG 1783  
QY  
481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
Db  
1784 CTGTCCAGGTGGCCCATCCCTTTATGGCTCCATTTGCTCAGCTCAGCTCAGCTGTC 1843  
QY  
501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
Db  
1844 ACTGCCTATGCTGCTGCGCGAGCGCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1903  
QY  
521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
Db  
1904 GTAGTATTGACAAGACGACTTGGCCCAATACTACAGC 1942  
QY  
541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553  
Db

## RESULT 10

AAB74800

ID AAB74800 standard; Protein; 553 AA.

XX AAB74800;

AC AAB74800;

XX 14-JUN-2001 (first entry)

XX Prostate tumour antigen predicted amino acid sequence for L1-12.

XX Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;

XX prostate cancer; immunogenic; cytostatic; vaccine.

XX Homo sapiens.

XX WO200125272-A2.

XX 12-APR-2001.

XX 04-OCT-2000; 2000WO-US27464.

XX 04-OCT-1999; 99US-0157455.

XX (CORI-) CORIXA CORP.

XX Xu J, Skeiky YAW, Reed SG, Cheever MA;

XX WPI; 2001-245062/25.

XX N-PSDB; AAH02530.

XX Prostate specific protein and its encoding polynucleotide, useful for

XX the treatment and diagnosis of prostate cancer -

XX Claim 3; Page 157-158; 276pp; English.

XX The present invention describes an isolated polypeptide (I) comprising

XX at least an immunogenic portion of a prostate tumour antigen protein or

XX its variant. (I) have cytostatic activity and can be used in vaccine

XX production. (I), prostate tumour antigen polynucleotides, an antigen

XX presenting cell (APC e.g. a dendritic cell) that expresses (I), and a

XX pharmaceutical composition containing (I) are useful for inhibiting the

XX development of cancer in a patient. Antibodies specific for prostate

XX specific proteins and oligonucleotides that hybridise to a

XX polynucleotide that encodes a prostate specific protein are useful

XX for detecting the presence or absence of a cancer or monitoring the

XX progression of a cancer, especially prostate cancer.

XX AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences

XX used in the exemplification of the present invention.

XX SQ Sequence 553 AA;

## Alignment Scores:

Pred. No.: 2,89e-208 Length: 553

Score: 2861.00 Matches: 553

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 44.58% Indels: 0

DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAB74800 (1-553)

QY 284 ATGGTCCAGAGGCTGTGGGTGAGCGCTGTGCGCACCGGAAAGCCAGCTCTTGTCTG 343

Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

QY 344 GTCAACCTGCTAACCTTTTGGCCCTGGAGGTGTCTTTTGGCGCAGGCATCACCTATGTGCGG 403

Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

QY 404 CCTCTGCTCTGGAAGTGGGGGTAGAGGAGAGATTTCATCACCATTGCTGCTGGCATTTGCT 463

Db 41 ProLeuLeuLeuGluValGlyValGlyGluGlyLysPheMetThrMetValLeuGlyIleGly 60



QY 464 CCAGTGTGGGCTGTGTGTCTCCGCTCTAGGCTCAGCAGTGCACCTGGCGTGA 523  
 DB 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CCCTATGGCCGCCGCCCTTCATCTGGGCACTGTCTTGGGATCCTCTGAGCCTC 583  
 DB 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCCAAGCGCGGTGGCTAGCAGGCTGTGTGCGCGATCCAGCGCCCTG 643  
 DB 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GAGCTGGACATGCTCATCTGGGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGT 703  
 DB 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCATGGAGGCCCTCTCTGACCTCTTCGCGGACCGGACCTGTGTGTGTGTGTGT 763  
 DB 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGTATGCTTCATGATCAGTCTTGGGGCTGTGTGTGTGTGTGTGTGTGTGT 823  
 DB 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGACACCGAGTGCCTTGGCCCTTACCTGGCAGCCAGGAGTGTCTTT 883  
 DB 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200  
 QY 884 GGCCTGTCTACCCCTCATCTTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 943  
 DB 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 CCAGCGCTGGGCCGCCCGGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1003  
 DB 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGCTGTCCATCCCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1063  
 DB 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACAGCTGTCTGCGGAGTCCCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1123  
 DB 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGGACATCATGCTTCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1183  
 DB 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
 QY 1184 TACCAGGCGGTGCCAGAGT 1243  
 DB 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
 QY 1244 GTTCGGATGGACCGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1303  
 DB 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTCATGACCGGT 1363  
 DB 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 QY 1364 GCTTTCCTGT 1423  
 DB 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380  
 QY 1424 TCAGCGCGCTCAGCGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483  
 DB 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
 QY 1484 TCCCTCTACACCGGAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1543  
 DB 401 SerLeuTyrHisArgGluGlyGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

QY 1544 GCTAGCAGTGGAGACGCTGTATGATCAGCAGCTTCTGCCAGCGCTTAAAGCTGGAGCTCCC 1603  
 DB 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 QY 1604 TTCCTAATGACACGCTGGT 1663  
 DB 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeu 460  
 QY 1664 TGGGGGCGCTCTGCTGT 1723  
 DB 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
 QY 1724 AGGTGTGTCTCGGCGCGGCGCATCTGTCTGACCTGTGTGTGTGTGTGTGTGTGTGTGT 1783  
 DB 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGTCCAGTGTGGCGCGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1843  
 DB 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTGCCCTATATGT 1903  
 DB 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleIleIleIleIleIle 540  
 QY 1904 GTAGTATTTGACAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1942  
 DB 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

## RESULT 11

ABB77575  
 ID ABB77575 standard; Protein; 553 AA.  
 XX ABB77575;  
 AC ABB77575;  
 DT 30-AUG-2002 (first entry)  
 DE Human mast cell related splice variant protein MC14 SEQ ID NO 13.  
 XX Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;  
 KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma.  
 OS Homo sapiens.  
 XX WO200246389-A2.  
 XX 13-JUN-2002.  
 XX 07-DEC-2001; 2001WO-US46180.  
 XX 08-DEC-2000; 2000US-251835P.  
 PR 14-MAR-2001; 2001US-275479P.  
 PR 28-MAR-2001; 2001US-279115P.  
 PR 02-APR-2001; 2001US-280143P.  
 XX (UNIO ) UCB SA.  
 XX Nocka K, Pirozzi G, Einstein R;  
 XX WPI; 2002-508560/54.  
 DR N-PSDB; ABN81324.  
 XX Novel isolated nucleic acids that are differentially expressed in mast  
 PT cells in patients with allergic hypersensitivity, encoding proteins  
 PT associated with mast cell regranulation and allergic hypersensitivity  
 PT  
 PS Claim 31; Page 117-119; 119pp; English.  
 XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),  
 CC corresponding to genes differentially expressed in mast cells following  
 CC activation or in patients with allergic hypersensitivity disease, (I)  
 CC that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of  
 CC (II) if at least 6 amino acids. (II) is useful for identifying binding

CC partners. (I) or (II) is useful for diagnosing or treating a disease  
 CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,  
 CC urticaria or atopic dermatitis or mastocytosis) in a subject which  
 CC involves determining the level of expression of (I) or (II). A computer  
 CC system, comprising a database containing information identifying the  
 CC expression level in a tissue or at least one mast cell of (I), is useful  
 CC for presenting information to identify the relative expression level of  
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic  
 CC response in a patient. The protein can also serve as a target that  
 CC modulate gene expression or activity and as an antigen to raise  
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying  
 CC agents that modulate expression of the protein or agents, such as  
 CC agonists or antagonists. The agonists or antagonists are useful for  
 CC modulating biological activity and function of (II) and thus are useful  
 CC for alleviating disease conditions such as allergic hypersensitivity,  
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.

XX Sequence 553 AA;

## Alignment Scores:

Pred. No.: 2,89e-208 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 23 Gaps: 0

US-09-759-143-110 (1-3410) x AB677575 (1-553)

QY 284 ATGGTCAGAGCTGGGTGAGCCGCTGTCGGCAGCCGAGAAAGCCAGCTCTTGCTG 343  
 Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
 QY 344 GTCAACCTGTAACCTTTGGCTGGAGGTGTGTTGGCCGAGGATCACCTATGTGCGC 403  
 Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40  
 QY 404 CCTCTGCTGCTGAAGTGGGGGTAGAGAGAGTTTCATGACCATGGTCTGGGCATTGGT 463  
 Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
 QY 464 CCAAGTCTGGGCTGTGTGTGTCGCGCTCCAGGCTCAGCCAGTGCACCATGGCGTGA 523  
 Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CGCTATGGCCGGCCGGCCCTTCATCTGGGCACTGTCTCTGGGCATCTGCTGAGCCTC 583  
 Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGCTGTCGCCGATCCCAAGGCCCTG 643  
 Db 101 PheLeuLeuProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GAGCTGGCACTGCTCATCTGCGGCTGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703  
 Db 121 GluLeuAlaLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGGAGCCCTGCTCTGACCTTCCTCCGGGACCCGACCATGTGCGCAGGCC 763  
 Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGTCTATGCTTCATGATCAGTCTTGGGGGCTGCTGGGCTACCTCTGCTGCC 823  
 Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGACACCACTGCTGCTGGCCCTTACCTGGGACCCAGGAGGAGTGCCTCTTT 883  
 Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200  
 QY 884 GGCTGCTACCTCATCTTCCTCAGCTGCTGGTAGGACCCACACTGCTGTGCTGAGGAG 943  
 Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

## RESULT 12

ABG61900

ID ABG61900 standard; Protein: 553 AA.

XX

AC ABG61900;

QY 944 GCAGCGCTGGGCCCCCAGCCAGCCAGCAGAGAGGCGCTGCGGCCCTCTCTTGTGCCCCAC 1003  
 Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGGTGTCCATGCCGGCGCCCTTGGCTTCCGGAACCTGGGCGCCCTGCTTCCCGGCTG 1063  
 Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACAGCTGTGCTCGCCGATGCCCCAGCCCTCGCCGGCTCTTCGTGCTGAGCTGTGC 1123  
 Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGCACATCATGACCTTCAGCTGCTTTTACACGGATTTCTGGCGAGGGCTG 1183  
 Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
 QY 1184 TACCAGCGCTGCGCCAGCTGAGCGGGCAGCCAGCGCCGAGCCGAGACACTATGATGAAGC 1243  
 Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
 QY 1244 GTTCGGATGGCAGCCCTGGGCTGTTCCTGCACTGGCCCATCTCCCTGCTTCTCTCTG 1303  
 Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTCATGCAACCGCTGGTGCAGCGATTTCGGCACCTCGAGCAGTCTATTGGCCAGCTGTGGA 1363  
 Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 QY 1364 GCTTTCCTGCTGGCTGCCGTCACATGCTGTGCCACAGTGTGCCGCTGGTGACAGCT 1423  
 Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380  
 QY 1424 TCAGCCCGCTCACCAGGTTTCACTTCTCAGCCCTGCAGATCTCCCTACACACTGGCC 1483  
 Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
 QY 1484 TCCCTCTACCAACCGGAGAGAGAGTGTCTCTGCCAAATACCGAGGGGACACTGGAGGT 1543  
 Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
 QY 1544 GCTAGCAGTAGGACACACCTGTATGACCAAGCTTCTGCCAGGCGCTTAGCCTGGAGTCCC 1603  
 Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 QY 1604 TTCCCTAATGGACACGTGGGTGTGGAGGAGTGGCTGCTCCACCTCCACCCCGCCTC 1663  
 Db 441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 460  
 QY 1664 TGGCGGCGCTCTGCTGCTGTATGCTTCCGTACGTGTGTGTGGTGTGGTGTGGAGGCC 1723  
 Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValGlyGluProThrGluAla 480  
 QY 1724 AGGCTGTTCCGGCGCGGCGCATCTGCTGGACCTCGCCATCTCCGATAGTGTCTCTCTG 1783  
 Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGTCCCAAGTGGCGCCCTCTCTGTTATGGGCTCCATTTGCCAGCTCAGCCAGTCTGTC 1843  
 Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTGCTATATGTTGTGTGTGCGCAGGCGTGGGTCTGGTGTGGCCATTACTTTGCTACACAG 1903  
 Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTGGACAGAGCGACTTGGCCCAATACTCAGCG 1942  
 Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

XX 15-AUG-2002 (first entry)  
 XX Prostate cancer-associated protein #101.  
 XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.  
 KW Mammalia.  
 OS WO200230268-A2.  
 XX PN 18-APR-2002.  
 XX PD 12-OCT-2001; 2001WO-US32045.  
 XX PF 13-OCT-2000; 2000US-0687576.  
 XX PR 08-DEC-2000; 2000US-0733288.  
 XX PR 08-DEC-2000; 2000US-0733742.  
 XX PR 24-JAN-2001; 2001US-263957P.  
 XX PR 16-MAR-2001; 2001US-276791P.  
 XX PR 16-MAR-2001; 2001US-276888P.  
 XX PR 06-APR-2001; 2001US-281922P.  
 XX PR 24-APR-2001; 2001US-286214P.  
 XX PR 30-APR-2001; 2001US-0847046.  
 XX PR 04-MAY-2001; 2001US-288589P.  
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX PI Gish KC, Mack DH, Willson KE, Afar D, Hevezi P;  
 XX WPI: 2002-471335/50.  
 XX DR N-PSDB; ABK92217.  
 XX PT Detecting a prostate cancer-associated transcript in a cell in a  
 PT patient, useful for diagnosing prostate cancer (PC) or screening  
 PT modulators of PC, by determining if prostate cancer-associated genes  
 PT are expressed in a prostate tissue  
 XX Claim 27; Page 386; 436pp; English.  
 XX The present invention relates to methods of detecting a prostate  
 CC cancer-associated transcript in a cell from a patient. The method  
 CC comprises contacting a biological sample from the patient with  
 CC prostate cancer-associated polynucleotides (designated PC genes) that  
 CC selectively hybridise to a sequence that is at least 80% identical  
 CC to them. The prostate cancer-associated polynucleotide sequences  
 CC are differentially expressed in prostate tumour tissue or in  
 CC prostate cancer and are derived from the tissues of various  
 CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).  
 CC The methods of the invention are useful for diagnosing and treating  
 CC prostate cancer in mammals. The prostate cancer-associated genes are  
 CC useful for diagnosing or treating prostate cancer, as well as for  
 CC identifying modulators of prostate cancer or agents that inhibit  
 CC prostate cancer. The nucleic acid sequences are particularly useful  
 CC in gene therapy, as a vaccine or in antisense applications.  
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins.  
 XX Sequence 553 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2,89e-208 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-759-143-110 (1-3410) x ABG61900 (1-553)  
 QY 284 ATGGTCCAGAGGCTGTGGGTGAGCGGCTGTCTGGCGGACCCGGAAGCCAGCTTCTGCTG 343  
 DB 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20

QY 344 GTCACCTGCTAACCTTTGGCCTGGAGGTGTGTTGGCCGAGCATCACCTATGTGCG 403  
 DB 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyValPro 40  
 QY 404 CCTCTGCTGCTGGAGTGGGTAGAGGAGTTCATGACCATGGTCTGGGCGATGGT 463  
 DB 41 ProLeuLeuGluValGlyValGluGlyPheMetThrMetValLeuGlyIleGly 60  
 QY 464 CCAGTGTGGGCTGTGTCTGTCTCCCGCTCAGGCTCAGCAGTGCAGCTGCGGTGA 523  
 DB 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CGCTATGGCCGCCGCCCTTTCATCTGGGCACTGTCTGGGCACTCTGCTGAGCCTC 583  
 DB 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCCAAGGCGGCTGTGTCTGAGCGGTGTGTGCGGAGTCCAGGCCCTG 643  
 DB 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GAGCTGGCACTGCTCATCTCTGGGCGTGGGCTGTGTGAGCTTCTGTGGCCAGGTGTCT 703  
 DB 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGGAGGCGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCCCGAGGCC 763  
 DB 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGTATGCTTTCATGATCAGTCTTGGGGCTGTCTGGGCTACCTCTCCCTGCGC 823  
 DB 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTCACTGGGACACCACTGCTGGCCCTGCTGGCCCTTCTGTGAGGAGTGTCTTT 883  
 DB 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGlyCysLeuPhe 200  
 QY 884 GGCTGTCTACCTCTCTCTCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 943  
 DB 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 GCAGCGTGGGCGCCCGGAGCCAGCAGAGGGCTGTGTGGCCCTCTCTTGTGCGCCAC 1003  
 DB 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGCTGTCCATGCGCGGCGCTTGGCTTCCGGAACCTGGCGGCTGTCTCTCCCGGCTG 1063  
 DB 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACAGCTGTGTGCGCGCATGCCCCCGACCTGCGCGGCTCTCTGTGTGTGTGTGTGTG 1123  
 DB 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGCACTCATGACCTTTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1183  
 DB 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
 QY 1184 TACAGGCGCTGCGCGAGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1243  
 DB 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
 QY 1244 GTTCGGATGGGCGAGCTGGGGCTGTCTGTGAGTGGCGCATCTCCCTGTGTGTGTGTGTGT 1303  
 DB 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTCATGGACCGGCTGT 1363  
 DB 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 QY 1364 GCTTTCCCTGT 1423  
 DB 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380  
 QY 1424 TCAGCGCGCTCACCGGCTTCACTTCTCAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1483

```

Db      381  SerAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400
Qy      1484 TCCTCTACACCGGAGAGAGAGTTCCTGCCCAATACCGAGGGGACACTGGAGGT 1543
Db      401  SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420
Qy      1544  GCTAGCAGTGAGGACAGCTGATGACACGCTTCCTGCCAGGCGCTAAGCGTGGAGCTCCC 1603
Db      421  AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440
Qy      1604  TTCCTCAATGACACGCTGGTGCTGGAGCAGTGGCGCTCTCCACCTCCACCGCGGTC 1663
Db      441  PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 460
Qy      1664  TGGGGGCTCTCCCTGTGTATGCTCTCCCTACGTGTGTGGTGGTGAGCCGACCGAGGCC 1723
Db      461  CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480
Qy      1724  AGGTGTGTTCCGGCGGGGCGCATCTGCTGGAGCTCGCCATCTCGGATAGTCCCTTCCTG 1783
Db      481  ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500
Qy      1784  CTGTCCAGGTGGCGCCATCCCTCTTATGGGCTCCATGTCAGCTCAGCCAGTCTGC 1843
Db      501  LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520
Qy      1844  ACTGCTTATATGTTGCTGCTCGGAGGCTGGGTCTGTGGCCATTTACTTTGCTACACAG 1903
Db      521  ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540
Qy      1904  GTAGTATTGACAGACGACTGGCGCAATATCTACGG 1942
Db      541  ValValPheAspLysSerAspLeuAlaIleTyrSerAla 553

```

## RESULT 13

ABB95222

ID ABB95222 standard; Protein: 553 AA.

AC ABB95222;

19-JUL-2002 (first entry)

Human L1-12 protein SEQ ID NO 113.

Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
gene therapy.

OS Homo sapiens.

PX US2002022248-A1.

PD 21-FEB-2002.

PF 12-JAN-2001; 2001US-0759143.

PR 25-FEB-1997; 97US-0806099.

PR 01-AUG-1997; 97US-0904804.

PR 09-FEB-1998; 98US-0020956.

PR 25-FEB-1998; 98US-0030607.

PR 14-JUL-1998; 98US-0115453.

PR 23-SEP-1998; 98US-0159812.

PR 15-JAN-1999; 98US-0232149.

PR 09-APR-1999; 99US-0288946.

PR 13-JUL-1999; 99US-0352616.

PR 12-NOV-1999; 99US-0439313.

PR 18-NOV-1999; 99US-0443686.

PR 14-JAN-2000; 2000US-0483672.

PR 27-MAR-2000; 2000US-0536857.

```

PR      10-AUG-2000; 2000US-0636215.
PR      29-AUG-2000; 2000US-0651236.
PR      06-SEP-2000; 2000US-0657279.
PR      02-OCT-2000; 2000US-0679426.
PR      10-OCT-2000; 2000US-0685166.
XX
PA      (XUJJ/) XU J.
PA      (DILL/) DILLON D C.
PA      (MITC/) MITCHAM J L.
PA      (HARL/) HARLOCKER S L.
PA      (JIAN/) JIANG Y.
PA      (KALO/) KALOS M D.
PA      (FANG/) FANGER G R.
PA      (RETT/) RETTER M W.
PA      (STOL/) STOLK J A.
PA      (DAYC/) DAY C H.
PA      (VEDV/) VEDVICK T S.
PA      (CART/) CARTER D.
PA      (LISX/) LI S X.
PA      (WANG/) WANG A.
PA      (SKEI/) SKEIKY Y A W.
PA      (HEPL/) HEPLER W T.
PA      (HEND/) HENDERSON R A.
XX
XX      Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI      Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI      Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
XX      WPI; 2002-255649/30.
XX
XX      New prostate-specific polynucleotides for diagnosing and treating
PT      diseases, in particular prostate cancer, and as markers for the
PT      progression of cancer.
XX
XX      Claim 2: SEQ ID NO 113; 87pp; English.
XX
XX      The present invention provides prostate-specific coding sequences and
CC      their encoded proteins. These can be used in the diagnosis and treatment
CC      of cancers, particularly prostate cancer. The present sequence is a
CC      protein described in the invention.
XX
XX      SQ Sequence 553 AA;

```

## Alignment Scores:

```

Pred. No.: 2,89e-208 Length: 553
Score: 2861.00 Matches: 553
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.58% Indels: 0
DB: 23 Gaps: 0

```

US-09-759-143-110 (1-3410) x ABB95222 (1-553)

```

Qy      284  ATGGTCCAGAGCTGTGGGTGAGCGCGCTGTCGGCAGCCGAAAGCCAGCTCTTGCTG 343
Db      1  MetValGlnArgLeuThrValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20
Qy      344  GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCCGCAGGCATCACCTATGTGCGG 403
Db      21  ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40
Qy      404  CCTCTGCTGCTGAAGTGGGGGTAGAGAGAGTTTCATGACCATGGTCTGGGCGATTGGT 463
Db      41  ProLeuLeuLeuValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60
Qy      464  CCAGTGTGGGCTGTGCTGTGTCGCCGCTCCCTAGGCTCAGCGAGTGCACCATGGGCGTGA 523
Db      61  ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisThrIleArgGly 80
Qy      524  CGCTATGGCGCGCGCGCGCTTTCATCTGGGCACTGCTCTGGGCACTCTCTGTCGATCCTG 583
Db      81  ArgTyrGlyArgArgProPheIleTyrAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

```

584 TTTCTCATCCAGGCGGCTGGCTAGCAGGCTGCTGTCGCCGATCCAGCCCTG 643  
 Db  
 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GAGCTGGCACTGCTATCTGCTGGCGTGGGCTGGCTGGACTTCTGTGGCAGGTGCTTC 703  
 Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGAGGCGCTGCTCTGACCTCTCCGGGACCCGACCACTGCTGCCAGGCC 763  
 Db 141 ThrProLeuGluAlaLeuLeuSerAspPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGTATGCTTCATGATCAGTCTTGGGGCTGCTGGGTACCTCTGCTGCTGCC 823  
 Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883  
 Db 181 IleAspTrpAspPheSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200  
 QY 884 GGCCTGCTCACCTCATCTTCTCCTCAGCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAG 943  
 Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 GCAGCGCTGGGCGCCACCGAGCCAGCAGCAGAAGGGCTGTGGGCGCTCTCTGTCGCCAC 1003  
 Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGTCTGCTCATGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063  
 Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123  
 Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGGCACTCATGACCTTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1183  
 Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
 QY 1184 TACCAGGGGCTGCCAGAGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1243  
 Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
 QY 1244 GTTCGATGGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1303  
 Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTTCATGGACGGCTGGCGAGGATTCGGCACTCGAGCAGTCTATTGGCCAGTGTGGCA 1363  
 Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 QY 1364 GCTTTCCTGCTGGCTGCCGCTGCCACATGCTGCTGCCAGTGTGGCGCTGCTGCTGCTGCT 1423  
 Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380  
 QY 1424 TCAGCGCGCTTCACCGGTTACCTTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1483  
 Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
 QY 1484 TCCCTCTACCGCGGAGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1543  
 Db 401 SerLeuTyrHisArgGluGlyGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
 QY 1544 GCTAGCAGTGGAGCAGCTGATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603  
 Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 QY 1604 TTCCTATGACAGCTGGTGGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1663  
 Db 441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460  
 QY 1664 TGCGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1723

Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
 QY 1724 AGGTTGGTTCGGGCGCGGCGCATCTGCTGACCTCGCATCTCGCATCTCGCATCTCGCTG 1783  
 Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGTCCAGGTGGGCGCCCATCTGCTGTTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1843  
 Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTGCTATATGGTGTCTGCGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1903  
 Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTGACAGAGCGACTGCTGCGCAAACTACTCAGCG 1942  
 Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

## RESULT 14

AAU10324  
 ID AAU10324 standard; Protein: 553 AA.  
 XX  
 AC AAU10324;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human PROST 03.  
 XX  
 XX Human; PROST 03; metastasis; prostate cancer; tumour; Immune response;  
 KW cytostatic; vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200181577-A2.  
 PD  
 XX 01-NOV-2001.  
 XX  
 XX 26-APR-2001; 2001WO-US13323.  
 PR 27-APR-2000; 2000US-200065P.  
 PR 20-APR-2001; 2001US-0200065.  
 XX

(SCHD ) SCHERING AG.

Lau T. Lin RJ, Parkes D, Parry G, Schneider DW, Steinbrecher R;  
 Van Heut PT, Wu J;

WPI; 2002-041404/05.

N-PSDB; AAS14962.

Novel PROST 03 polypeptides and polynucleotides useful in research,  
 diagnosis and therapeutic applications, particularly for use in cancer  
 therapeutics

Claim 14; Fig 2; 77pp; English.

The invention relates to an isolated PROST 03 polypeptide (I) and to  
 the polynucleotide (II) encoding PROST 03. Fragments of (I) were used  
 to generate antibodies (III) to PROST 03. (III) is useful for selectively  
 destroying a cell expressing (I), and for treating a disease-state  
 associated with expression of PROST 03 in a human patient. (III) is  
 useful for diagnosing metastasis associated with (I), in a subject.  
 (I) is also useful for diagnosing and treating diseases of cell  
 proliferation such as prostate cancers. (I) is also useful for generating  
 antibodies to PROST 03. (III) is useful in detecting the levels of PROST  
 03 polypeptides in cells and tissues, and in targeting drugs to primary  
 and metastatic tumours. (I) is also useful for stimulating immune  
 response to PROST 03 containing cells. (II) is useful in diagnostic  
 assays for detecting the levels of polynucleotides encoding PROST 03 in  
 cells and tissues. (II) is useful as DNA probes, as targets for antisense  
 and ribozyme therapy, and as templates for the production of antisense  
 polynucleotides. (I) and (II) are useful in research, biological,

CC clinical and therapeutic purposes. The present sequence represents  
 CC the amino acid sequence of human PROST 03.

XX Sequence 553 AA;

#### Alignment Scores:

Pred. No.: 2,89e-208 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 23 Gaps: 0

US-09-759-143-110 (1-3410) x AAU10324 (1-553)

QY	284	ATGTCCTCAGAGCTGTGGGTGAGCCGCTGCTGGCGGACCGGAAAGCCAGCTTGTGCTG	343
Db	1	MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu	20
QY	344	GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTGGCGGAGGATCACCTATGTGCGG	403
Db	21	ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyValPro	40
QY	404	CCTTCTGCTGGAAGTGGGGTAGAGAGAGTTTCATGACCATGGTCTGGGCAATTGGT	463
Db	41	ProLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly	60
QY	464	CAAGTCTGGGCTGTGCTGCTGCCGCTCTAGGCTCAGCCAGTGACACTGGGCTGGA	523
Db	61	ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly	80
QY	524	CGCTATGGCGCGCGGCGCTTCTATCTGTGGGCACTGTCTTGGGCACTCTGCTGAGCCTC	583
Db	81	ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu	100
QY	584	TTTCTCATCCAGGCGCGCTGGCTAGCAGGCTGCTGCGCCGAGATCCAGGCGCCCTG	643
Db	101	PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu	120
QY	644	GAGCTGGCACTGTCTATCTTGGCGCTGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC	703
Db	121	GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe	140
QY	704	ACTTCAGTGGAGCCCTGTCTGTGACTTTCGCGGAGCCGAGCCAGTGTGCGCCAGGCC	763
Db	141	ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla	160
QY	764	TACTCTGTCTATGCTTATGATCATCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCC	823
Db	161	TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla	180
QY	824	ATTGACTGGGACACCACTGCTTCTCCTCAGCTGGTAGCAGCCACACTGCTGTGCTCAGGAG	943
Db	201	GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu	220
QY	944	GCAGCGTGGCGCCCGCCAGGAGGCTGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG	1003
Db	221	AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis	240
QY	1004	TGCTGTCCATGCGGCGCGCTTGTGCTTTCGGAACCTGGGCGCCCTGCTTCCCGCGCTG	1063
Db	241	CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu	260
QY	1064	CACCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1123
Db	261	HisGlnLeuCysCysArgMetProArgThrLeuArgLeuPheValAlaGluLeuCys	280
QY	1124	AGCTGGATGGCACTCATGACTTTCAGCTGTTTACACGATTTTGTGGCGGAGGCTG	1183

Db	281	SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu	300
QY	1184	TACCAGGCGTGCAGAGCTGAGCGGCGGAGCCAGCCAGGAGACACTATATGAGAGC	1243
Db	301	TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly	320
QY	1244	GTTCCGATGGCAGCCTGGGCTGTTCCTGCACTGCGCCATCTCCCTGCTTCTCTCTG	1303
Db	321	ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu	340
QY	1304	GTCAATGACCGGCTGTGTCAGCGATTCGGCACTCGAGCTATTTGTCAGCTGTGGCA	1363
Db	341	ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla	360
QY	1364	GCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1423
Db	361	AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla	380
QY	1424	TCAGCCCGCTCAGCGGTTTCACTTCTCAGCCCTGAGATCTCCCTGCTGCTGCTGCTG	1483
Db	381	SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla	400
QY	1484	TCCCTCTACCAACCGGAGAGAGGCTTCTCCTCCCAATACCGAGGAGACACTGGAGGT	1543
Db	401	SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly	420
QY	1544	GCTAGCAGTGGAGACGCTGTGATGACAGCTTCTCCTCCAGGCGCTAAGCCTGGAGTCCC	1603
Db	421	AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro	440
QY	1604	TTCCCTATGACACGCTGGTGGTGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1663
Db	441	PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu	460
QY	1664	TGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1723
Db	461	CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla	480
QY	1724	AGGTGGTTCGGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1783
Db	481	ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu	500
QY	1784	CTGTCGCGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1843
Db	501	LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal	520
QY	1844	ACTCCCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1903
Db	521	ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln	540
QY	1904	GTAGTATTGACAGAGCGGCTTGGCCCAATACTCAGCG	1942
Db	541	ValValPheAspLysSerAspLeuAlaLysTyrSerAla	553

#### RESULT 15

AAU82643

ID AU82643 standard; Peptide: 553 AA.

XX AU82643;

XX 23-APR-2002 (first entry)

XX Human breast tumour polypeptide clone #3.

XX Human; breast tumour polypeptide; breast cancer; cytostatic;  
 immunostimulant.

XX Homo sapiens.

XX WO200198339-A2.

XX 27-DEC-2001.

XX



PF 12-JUN-2001; 2001WO-US19032.  
XX  
PR 22-JUN-2000; 2000US-0602877.  
PR 12-OCT-2000; 2000US-0687507.  
PR 06-FEB-2001; 2001US-0778381.  
XX  
XX (CORI-) CORIXA CORP.  
PI Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;  
XX  
XX WPI; 2002-147792/19.  
DR N-PSDB; ABR29012.  
XX  
XX Polynucleotides encoding breast tumour polypeptides, useful for treating  
PT breast cancer or stimulating an immune response  
XX  
XX \*Claim 2; Page 144-145; 150pp; English.  
XX  
XX The invention relates to polynucleotides encoding breast tumour  
CC polypeptides. The sequences are useful for treating cancer, preferably  
CC breast cancer, in a patient or for stimulating an immune response. The  
CC polynucleotides and polypeptides are also useful in the diagnosis and  
CC monitoring of breast cancer. A method for detecting the presence of a  
CC cancer in a patient, comprises obtaining a biological sample from the  
CC patient, contacting the biological sample with a binding agent that binds  
CC to a breast tumour polypeptide, detecting in the sample an amount of  
CC polypeptide that binds to the binding agent, and comparing the amount of  
CC polypeptide to a predetermined cut-off value, therefore determining the  
CC presence of a cancer in the patient. Sequences AAU82641-AAU82655  
CC represent human breast tumour polypeptides of the invention.  
XX  
SQ Sequence 553 AA;

## Alignment Scores:

Pred. No.: 2,89e-208 Length: 553  
Score: 2861.00 Matches: 553  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 44.58% Indels: 0  
DB: 23 Gaps: 0

US-09-759-143-110 (1-3410) x AAU82643 (1-553)

QY 284 ATGGTCACAGGCTGTGGTGAGCCGCTGCTGGCAGCCAGGAGCCAGCTGTGCTG 343  
DB 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
QY 344 GTCAACCTGCTACCTTTGGCCTGGAGGTGTGTTGGCCGAGGATACCTATGTGCG 403  
DB 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTrpValPro 40  
QY 404 CCTCTGCTGTGAAGTGGGGGTAGAGGAGAGTTCATGACCATGCTGTGGGCAATTGGT 463  
DB 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
QY 464 CCAGTGTGGGCTGTGTGTGTGTCCTGCTAGGCTAGCCAGGATGACCACTGGGTGGA 523  
DB 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
QY 524 CGTATGGCCGCGCGGCCCTTCATCTGGGCACTGTCTTGGGCACTCTGCTGAGCCCTC 583  
DB 81 ArgTrpGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
QY 584 TTCTCTATCCCAAGGCGCGCTGGTAGGAGGCTGCTGTGCCGATCCAGGCCCTG 643  
DB 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
QY 644 GAGCTGCACCTGCTCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703  
DB 121 GluLeuAlaLeuLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
QY 704 ACTCCACTGAGGCGCTGTCTCTGACCTTCTCCGGGACCCGACCACTGTGCGCAGGCC 763  
DB

Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
QY 764 TACTCTGTCTATGCTTCATGATCAGTCTTTGGGGGCTGCTGGCTGCTGCTGCTGCTGCTG 823  
Db 161 TyrSerValTrpAlaPheMetIleSerLeuGlyGlyCysLeuGlyTrpLeuLeuProAla 180  
QY 824 ATTGACTGGGACACCACTGCTGGCCCTTACTGCTGGCACCAGGAGGAGTCCCTCTTT 883  
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTrpLeuGlyThrGlnGluGlnCysLeuPhe 200  
QY 884 GGCCTGCTCACCCTCATCTCTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943  
Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
QY 944 GCACGCTGGGCGCCACCAGGAGGAGGCTGCTGGCCCTCTCTGCTGCTGCTGCTGCTG 1003  
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
QY 1004 TGCTGTCCATGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1063  
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
QY 1064 CACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1123  
Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
QY 1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTACACGAGTTCGTTGGGCGAGGGCTG 1183  
Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheThrThrAspPheValGlyGluGlyLeu 300  
QY 1184 TACCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1243  
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
QY 1244 GTTCGGATGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1303  
Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
QY 1304 GTCATGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1363  
Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTrpLeuAlaSerValAla 360  
QY 1364 GCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423  
Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValAlaThrAla 380  
QY 1424 TCAGCGCGCTCACCGGCTTCACTTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1483  
Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTrpThrLeuAla 400  
QY 1484 TCCCTCTACCAACCGGAGAGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1543  
Db 401 SerLeuTrpHisArgGluGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
QY 1544 GCTACGAGTGGAGGAGGCTGATGACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1603  
Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
QY 1604 TTCCCTTAATGACAGCTGGGTGCTGGAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1663  
Db 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProProAlaLeu 460  
QY 1664 TGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1723  
Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
QY 1724 AGGTGGTTCGGGCGCGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1783  
Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
QY 1784 CTGTCCCGAGTGGCGCCCATCTCTGTTATGGCTCATTTGCTGCTGCTGCTGCTGCTGCTG 1843  
Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520



QY 1844 ACTGCCTATATGTTGCTGCGCAGGCGCTGGTCTGGTCCCAATTTACTTTGCTACACAG 1903  
 Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaTleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTTTACAAAGAGCAGCTTGGCCAAATACTCAGCG 1942  
 Db 541 ValValPheAspLysSerAlaAlaLysTyrSerAla 553

## RESULT 16

AAB74830  
 ID AAB74830 standard; Protein; 1079 AA.

XX AAB74830;

XX 14-JUN-2001 (first entry)

XX Prostate tumour antigen amino acid sequence for a fusion protein.

XX Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;  
 XX prostate cancer; immunogenic; cytostatic; vaccine.

XX Homo sapiens.

XX WO200125272-A2.

PN 12-APR-2001.

XX 04-OCT-2000; 2000WO-0257464.

XX 04-OCT-1999; 99US-0157455.

XX (CORI-) CORIXA CORP.

PI Xu J, Skelky YAW, Reed SG, Cheever MA;

XX WPI; 2001-245062/25.

XX Prostate specific protein and its encoding polynucleotide, useful for  
 PT the treatment and diagnosis of prostate cancer -

XX Disclosure; Page 272-276; 276pp; English.

CC The present invention describes an isolated polypeptide (I) comprising  
 CC at least an immunogenic portion of a prostate tumour antigen protein or  
 CC its variant. (I) have cytostatic activity and can be used in vaccine  
 CC production. (I), prostate tumour antigen polynucleotides, an antigen  
 CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a  
 CC pharmaceutical composition containing (I) are useful for inhibiting the  
 CC development of cancer in a patient. Antibodies specific for prostate  
 CC specific proteins and oligonucleotides that hybridize to a  
 CC polynucleotide that encodes a prostate specific protein are useful  
 CC for detecting the presence or absence of a cancer or monitoring the  
 CC progression the progression of a cancer, especially prostate cancer.  
 CC AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences  
 CC used in the exemplification of the present invention.

XX Sequence 1079 AA;

## Alignment Scores:

Pred. No.: 3,57e-208 Length: 1079  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAB74830 (1-1079)

QY 284 ATGGTCCAGAGGCTGGGTGAGCGCGCTGCTGGGACCGCAAGCCAGCTCTTGTG 343

Db 527 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 546

QY 344 GTCAACCTGCTAAACCTTTGGCGTGGAGGTTGTTTGGCGCAGGAGCATCACCTATGTGCGC 403  
 Db ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyrValPro 566  
 QY 404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGATTCATGACCATGGTGGTGGCAITGGT 463  
 Db ProLeuLeuGluValGlyValGluGlyPheMetThrMetValLeuGlyIleGly 586  
 QY 464 CCAGTGTGGCGCTGCTGCTGCTCCCGCTCCCTAGGCTCAGCAGTACCAGCTGGCGTGA 523  
 Db ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 606  
 QY 524 CGTATGGCGCGCGCGCTTTCATCTGGGCACTGCTTGGGCACTCTCTGGGCACTCTCT 583  
 Db ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 626  
 QY 584 TTTCTCATCCCAAGCGCGCTGGCTAGCAGGGGTGCTGTGCGCGGATCCAGGCCCTG 643  
 Db PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 646  
 QY 644 GAGCTGGCACTGCTCATCTGGGCGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 703  
 Db GluLeuAlaLeuLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 666  
 QY 704 ACTCCACTGGAGCGCTGCTCTCTGACCTCTCTCGGAGCCCGGACCACTGTGCGCAGGCC 763  
 Db ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysAspGlnAla 686  
 QY 764 TACTCTGTCTATGCTTCATGATCAGTCTGGGGCTGCTGGGCTGCTGGCTACCTCTG 823  
 Db TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 706  
 QY 824 ATTGACTGGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883  
 Db IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 726  
 QY 884 GGCTGCTCACCCTCATCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943  
 Db GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGlu 746  
 QY 944 GCAGCGTGGCG 1003  
 Db AlaAlaLeuGlyProThrGluProAlaGluLeuSerAlaProSerLeuSerProHis 766  
 QY 1004 TGCTGTCCATGCG 1063  
 Db CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 786  
 QY 1064 CACCAGCTGTGCTGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1123  
 Db HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 806  
 QY 1124 AGCTGGATGCACTCATGACCTTTCAGCTGCTTTTACACGAGTTCCTGGCGAGGGCTG 1183  
 Db SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluLeu 826  
 QY 1184 TACCAGGCGCTGCCAGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1243  
 Db TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTrpAspGluGly 846  
 QY 1244 GTTCGGATGGCAGCGCTGGCGCTGTTCTGCTGCTGCGCGCATCTCCCTGGTCTCTCTG 1303  
 Db ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 866  
 QY 1304 GTCATGACCGCGCTGGTGGAGCGATTTGGCAGTCTGCTATTTGGCCAGTGTGGCA 1363  
 Db ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 886  
 QY 1364 GCTTTCCCTGCTGCTGCG 1423  
 Db AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 906  
 QY 1424 TCAGCGCGCGCTCACCGCGGTTCACTTCTCAGCGCGCTGCGATCTGCGCTACACACTGG 1483



Qy	860	GGCACCCAGGAGGAGTGCCCTCTTTGGGCTGCTCACCCCTCATCTTCTCTCACTCGGTAGCA	919
Db	342	GlyThrGlnGluLucysLeuPheGlyLeuThrLeuIlePheLeuThrCysValAla	361
Qy	920	GCCACACTGCTGGTGTGAGGAGCAGCGCTGGGGCCCCCAGCGAGCAGAGAAGGGCTG	979
Db	362	AlaThrLeuLeuValAlaGluGluAlaLeuGlyProThrGluProAlaGluGlyLeu	381
Qy	980	TCGGCCCCCTCTGTGTCGCCCCACTGCTGTCATGCGGGCGGCTTGGCTTTCCCGAAC	1039
Db	382	SerAlaProSerLeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsn	401
Qy	1040	CTGGGGCCCTGCTTCCCGGCTCCACCAAGCTGTGCTGCCGATGCGCCCGCACCCCTCCGC	1099
Db	402	LeuGlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArg	421
Qy	1100	CGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTAC	1159
Db	422	ArgLeuPheValAlaGluLeuCysSerTrpMetAlaLeuMetThrPheThrLeuPheTyr	441
Qy	1160	ACGGATTTCTGGCGAGGGGCTGTACCAAGGGGTGCCAGAGCTGAGCGGGCACCAG	1219
Db	442	ThrAspPheValGlyGluGlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGlu	461
Qy	1220	GCCCGAGACACTATGATGAAGCGGTTCGATGGCAGCGCTGGGGTGTTCCTGCAGTGC	1279
Db	462	AlaArgArgHisTyrAspGluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCys	481
Qy	1280	GCCATCTCCCTGGTCTCTCTGGTTCATGGACGGGCTGGTGACAGCGATTTCGGCACTCGA	1339
Db	482	AlaIleSerLeuValPheSerLeuValMetAspArgLeuValGlnArgPheGlyThrArg	501
Qy	1340	GCAGCTATTTCGGCACGTGTGGCAGCTTTCCTCTGGCTCCGGTGCCACATGCCCTGCC	1399
Db	502	AlaValTyrLeuAlaSerValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSer	521
Qy	1400	CACAGTGTGGCGGTGTGACAGCTTCACCGCCCTCACCGGGTTCACCTTCTCAGCCCTG	1459
Db	522	HisSerValAlaValAlaThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeu	541
Qy	1460	CAGATCTGCCCTACACACTGGGCTCCCTCTACCACCGGAGAGGAGGTTCCTGCC	1519
Db	542	GlnIleLeuProTyrThrLeuAlaSerLeuTyrHisArgGlnGlyGlnValPheLeuPro	561
Qy	1520	AAATACCGAGGGACACTGGAGGTGCTAGCAGTGAGGACACCGCTGTATGACCAAGCTTCC	1579
Db	562	LysTyrArgGlyAspThrGlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeu	581
Qy	1580	CCAGGCCCTAAGCTTGAGCTCCCTTCCTTAATGGACACGTGGGTGCTGAGGACAGTGGC	1639
Db	582	ProGlyProLysProGlyAlaProPheProAsnGlyHisValGlyAlaGlyGlySerGly	601
Qy	1640	CTGCTCCCACTCCACCCGGGCTCTGCGGGCCCTCTGCCGTGTATGCTCCGACAGTGTG	1699
Db	602	LeuLeuProProProAlaLeuCysGlyAlaSerAlaCysAspValSerValArgVal	621
Qy	1700	GTGGTGGGTGAGCCACCGAGGGTGGTTCCGGGCGGGGCATCTGCCTGGACCTC	1759
Db	622	ValValGlyGluProThrGluAlaArgValValProGlyArgGlyIleCysLeuAspLeu	641
Qy	1760	GCCATCTCGTAGTGCCTTCCTGCTGCCAGGTGGCCCCCATCCCTGTTTATGGGGCTCC	1819
Db	642	AlaIleLeuAspSerAlaPheLeuLeuSerGlnValAlaProSerLeuPheMetGlySer	661
Qy	1820	ATTGTCCAGCTACGCCAGTGTGTCACTGCCCTATATGTTGTCTCCGCGAGCCCTGGGCTG	1879
Db	662	IleValGlnLeuSerGlnSerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeu	681
Qy	1880	GTCCCATTTTACTTTGCTACACAGTAGTATTTTGACAGAGCGACTTTGGCCAAATACTCA	1939
Db	682	ValAlaIleTyrPheAlaThrGlnValValPheAspLysSerAspLeuAlaLysTyrSer	701

Qy	1940 GCG 1942
Db	702 Ala 702
RESULT 18	
AAAM01318	
ID	AAAM01318 standard; Protein; 595 AA.
XX	AAAM01318;
AC	
XX	
DT	04-OCT-2001 (first entry)
XX	
DE	Alpha prepro-P501S recombinant protein amino acid sequence.
XX	
KW	Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW	cytostatic; gene therapy; metastasis.
XX	
OS	Homo sapiens.
XX	
PN	WO200151633-A2.
XX	
PD	19-JUL-2001.
XX	
PF	16-JAN-2001; 2001WO-US01574.
XX	
PR	14-JAN-2000; 2000US-0483672.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI	Kalos MD, Fanger GR, Day CH, Retter MW, Stoik JA, Skeiky YAW;
PI	Wang A, Meagher MJ;
XX	
DR	WPI; 2001-425873/45.
XX	
PT	New polynucleotide encoding a prostate-specific protein, for
PT	diagnosing, monitoring and treating prostate cancer in a patient and
PT	for use in vaccines -
XX	
PS	Example 17; Page 541-543; 543pp; English.

```

Db      81  LeuGluLysArgGluAlaGluAlaMetValLeuGlyIleGlyProValLeuGlyLeuVal 100
      482  TGTGTCCCGCTCTAGGCTCAGCAGTGTACCACTGGCGTGGAGCTATGGCCGCGCGG 541
Db      101  CysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTrpGlyArgArg 120
QY      542  CCTTCATCTGGCACTGCTCTGGGCATCTCTGGCATCTCTGGCATCTCTCTCTCTCTCT 601
Db      121  PropHeileTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIleProArgAla 140
QY      602  GCGTGGCTAGAGGCTGTGTGCGGATCCAGGATCCAGGCTGGAGCTGGAGCTGTCTATC 661
Db      141  GlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeuLeuIle 160
QY      662  CTGGGCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
Db      161  LeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGluAlaLeu 180
QY      722  CTCTCTGACCTCTTCGGGACCCGACCACTGTGCGGAGGCTACTCTGTCTATGCGCTC 781
Db      181  LeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTrpSerValTrpAlaPhe 200
QY      782  ATGATCAGCTTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
Db      201  MetIleSerLeuGlyGlyCysLeuGlyTrpLeuLeuProAlaIleAspTrpAspThrSer 220
QY      842  GGCCTGGCCCTTACCTGGGACCCAGGAGTGCCTCTTGGGCTGCTGCTGCTGCTGCTGCT 901
Db      221  AlaLeuAlaProTrpLeuGlyThrGlnGluCysLeuPheGlyLeuLeuThrLeuIle 240
QY      902  TTCCTCAGCTGCTGAGCAGCACACTGCTGGTGGTGGAGGAGGAGGAGGAGGAGGAGGAG 961
Db      241  PheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaLeuGlyProThr 260
QY      962  GAGCAGCAGAGAGGCTGCGGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
Db      261  GluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCysArgAla 280
QY      1022  CGCTTGGCTTCCGGAACCTGGGCGCCCTGCTTCCCGGCTGCACCACTGCTGCTGCTGCT 1081
Db      281  ArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArg 300
QY      1082  ATGCCCCGACCTCGCGCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
Db      301  MetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAlaLeuMet 320
QY      1142  ACCTTCAGCTGCTTTTACACGATTTCTGCGGCGAGGGCTGTACAGGCGGTGCCGAGA 1201
Db      321  ThrPheThrLeuPheTrpThrAspPheValGlyGluGlyLeuTrpGlnGlyValProArg 340
QY      1202  GCTGAGCGGGCAGCGCGGAGACACACTATGATGAAGCGTTCGGATGGGAGGCTG 1261
Db      341  AlaGluProGlyThrGluAlaArgArgHisTrpAspGluGlyValArgMetGlySerLeu 360
QY      1262  GSGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1321
Db      361  GlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMetAspArgLeuVal 380
QY      1322  CAGCGATTCGGCAGCTGCTATTTGGCGAGTGTGGCAGCTTTCCCTGCTGCTGCTGCT 1381
Db      381  GlnArgPheGlyThrArgAlaValTrpLeuAlaSerValAlaAlaPheProValAlaAla 400
QY      1382  GGTGCCACATGCTGTCCACAGTGTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1441
Db      401  GlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAlaAlaLeuThrGly 420
QY      1442  TTCACCTTCTCAGCCCTGCAGATCTGCGCTTACACAGTGGCTCTCTCTCTCTCTCTCT 1501
Db      421  PheThrPheSerAlaLeuGlnIleLeuProTrpThrLeuAlaSerLeuTrpHisArgGlu 440
QY      1502  AAGCAGGTGTCTCTCCCAATACCGAGGGGACACTGGAGGTGTGTGTGTGTGTGTGTGTGT 1561

```

## RESULT 19

AAM50662

ID AAM50662 standard; Protein: 530 AA.

XX AAM50662;

AC AAM50662;

XX 08-APR-2002 (first entry)

DE Thioedoxin-ubiquitin-P501S(aal-320)-His triple fusion protein.

KW Thioedoxin; ubiquitin; P501S; prostate; tumour; antigen; cancer;

KW vaccine; therapy; human.

XX Chimeric - Homo sapiens.

OS Chimeric - Escherichia coli.

XX Key Location/Qualifiers

FH Protein 1..109

FT /label= Thioedoxin

FT Peptide 110..125

FT /label= Linker

FT Protein 126..199

FT /label= Ubiquitin

FT Cleavage-site 201..202

FT /label= Ubiquitin

FT Misc-difference 202..203

FT /note= "translation of the DNA sequence of Figure

12a (ABA91284) places a Val residue between

Met-202 and Gln-203"

FT Protein 203..521

FT /label= P501S

FT Misc-difference 208

FT /note= "Val at position 208 is not encoded

by the DNA sequence of Figure 12a

FT Peptide 522..530

FT /label= Histidine\_tail

FT WO200200892-Al.

XX 03-JAN-2002.

XX 19-JUN-2001; 2001WO-EP06952.



Novel isolated nucleic acids that are differentially expressed in mast cells in patients with allergic hypersensitivity, encoding proteins associated with mast cell degranulation and allergic hypersensitivity

Claim 31: Page 103-104; 119pp; English.

The invention relates to isolated nucleic acid (ABN81319-ABN81324), corresponding to genes differentially expressed in mast cells following activation or in patients with allergic hypersensitivity disease, (I) that encodes proteins (ABN77569-ABN77575) (II) or a protein fragment of (II) if at least 6 amino acids. (II) is useful for identifying binding partners. (I) or (II) is useful for diagnosing or treating a disease state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma, urticaria or atopic dermatitis or mastocytosis) in a subject which involves determining the level of expression of (I) or (II). A computer system, comprising a database containing information identifying the expression level in a tissue or at least one mast cell of (I), is useful for presenting information to identify the relative expression level of (I). (II) is used as a marker to detect, diagnose or identify an allergic response in a patient. The protein can also serve as a target that modulate gene expression or activity and as an antigen to raise polyclonal or monoclonal antibodies. (II) is useful for identifying agents that modulate expression of the protein or agents, such as agonists or antagonists. The agonists or antagonists are useful for modulating biological activity and function of (II) and thus are useful for alleviating disease conditions such as allergic hypersensitivity, seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.

Sequence 326 AA;

Alignment Scores:

Pred. No.: 6,64e-107 Length: 326  
Score: 1522.50 Matches: 300  
Percent Similarity: 58.85% Conservative: 6  
Best Local Similarity: 57.69% Mismatches: 12  
Query Match: 23.72% Indels: 202  
DB: 23 Gaps: 2

US-09-759-143-110 (1-3410) x ABN77571 (1-326)

QY	347	AACTCTGCTACCTTTGGCTGGAGGTGTTGGCCGAGGAGCATCACCTATGTGCGGCT	406
DB	9	SerLeuValProLeuProLeuAlaLeuTyfLeuSerGlnProLeuThrThrSer	28
QY	407	CTCTGCTGGAAGTGGGGTAGAGGAGAAGTTTCATGACCATGCTGCTGGGCATTGGTCCA	466
DB	29	LeuLeu-----AlaGlyIleGlyPro	35
QY	467	GTGCTGGGCTGTGTCTGCTCCGCTCTAGGCTCAGCCAGTGACCATGGCGTGGACGC	526
DB	36	ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg	55
QY	527	TATGGCCGGCGCGCCCTTCATCTGGGCACTGTCTCTGGGCATCTCTGCTGAGCTCTTT	586
DB	56	TyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe	75
QY	587	CTCATCCCAAGGCGGCTGGCTAGCAGGCTGTGTGCCCGATCCAGGCCCTGGAG	646
DB	76	LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu	95
QY	647	CTGGCACTGCTATCTTGGCGCTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTCACT	706
DB	96	LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr	115
QY	707	CCACTGGAGCCCTGCTCTCTGACCTCTCCGGGACCCGAGCAGTGTCCGAGCGCTAC	766
DB	116	ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr	135
QY	767	TCNCTATGCTTCATGATCATCTTGGGGCTGCTGGGTACCTCTGCTGCTGCCATT	826
DB	136	SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle	155

QY	827	GACTGGACACCACTGCTCCCTGCGCCCTACTCTGGCACCAGGAGGAGTCTTCTTGGC	886
DB	156	AspTrpAspThrSerAlaLeuAlaProTyfLeuGlyThrGlnGluGluCysLeuPheGly	175
QY	887	CTGCTCACCCCTCTCTCTCACTGCTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA	946
DB	176	LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla	195
QY	947	GCCTGGGCCCCACCCGAGCCAGAGGGCTGTGCGCCCTCTCTTGTGCCCCCACTGC	1006
DB	196	AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys	215
QY	1007	TGTCATGCGCGGCGCCCTTGGCTTCCGGAACCTGGCGCCCTGTCTCCCGGCTGCAC	1066
DB	216	CysProCys-----	218
QY	1067	CAGTGTGCTGCCCATGCCCGCCACCCCTGCGCGGGCTCTTGTGGCTGAGTGTGCAGC	1126
DB	218	-----	218
QY	1127	TGGATGGCACTCATGACCTTCACGCTGTTTACACGGATTTCGTGGCGGAGGCTGTAC	1186
DB	218	-----	218
QY	1187	CAGGCGGTGCCAGAGTGTGAGCGGCGGACCGGAGCCCGGAGACACTATGATGAAGCGTT	1246
DB	218	-----	218
QY	1247	CGGATGGCAGCCTGGGGCTGTTCTCTGACGTGCGCCATCTCCCTGGTCTTCTCTCGTC	1306
DB	218	-----	218
QY	1307	ATGGACCGGCTGTGTGTCAGCGATTGCGCACTCGAGCAGTCTATTATTTGGCAGTGTGCAGCT	1366
DB	218	-----	218
QY	1367	TTCCCTGTGGCTGCCGGTGCCACATGCTGTCCACAGTGTGGCGGTGGTGCACAGCTTCA	1426
DB	218	-----	218
QY	1427	GCGGCCCTCACCGGGTTACCTTCTCAGCCCTGCAGATCTCTGCCCTACACACTGGCCTCC	1486
DB	218	-----	218
QY	1487	CTCTACCACCGGAGAGAGAGTGTCTCTGCCAAATACCGAGGGGACACTGGAGGTCT	1546
DB	218	-----	218
QY	1547	AGCAGTAGGACAGCCTGATGACAGCTTCTCTGCCAGGCCCTTAAGCTGGAGTCCCTTC	1606
DB	219	-----Gln-AlaLeuSerLeuGluLeuProSe	227
QY	1607	CCTAATGGACACGTGGGTGTGGAGCAGTGGCTGTCCACCTCCACCCGCGCTCTGC	1666
DB	227	rLeuMetAspThrTrpValLeuGluAlaValAlaCysSerHisLeuHisProArgSerAl	247
QY	1667	GGGGCTCTGCTGTGTGTCTCCGTACGTGTGGTGGTGGTGGAGCCACCGAGGCCAGG	1726
DB	247	aglyProLeuProValMetSerProTyfValTrpTrpValSerProProArgProG	267
QY	1727	GTGGTTCGGCGCGGGGATCTGCTGACCTCGGCATCTCGGATCTCGGATGCTCTCTGCTG	1786
DB	267	yTrpPheArgAlaGlyAlaSerAlaTrpThrSerProSerTrpIleValProSerCysC	287
QY	1787	TCCAGGTGGCCCATCTCTGTTTATGGCTTCCAGCTTGTCCAGCTCAGCAGCTGTCACT	1846
DB	287	sProArgTrpProHisProCysLeuTrpAlaProLeuSerSerAlaSerLeuSerLe	307
QY	1847	GCCTATATGTGTCTGCGCGAGGCTGGGTCTGCTGCTGCTGCTGCTTGTGTACACAGG	1904
DB	307	uProIleTrpCysLeuProGlnAlaTrpValTrpSerProPheThrLeuLeuHisArg	326





DE P553S splice variant P553S-14 amino acid #2.  
XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
KW cyrostatic; gene therapy; metastasis.  
XX  
OS Homo sapiens.  
XX  
PN WO200151633-A2.  
XX  
XX 19-JUL-2001.  
XX  
PF 16-JAN-2001; 2001WO-US01574.  
XX  
XX 14-JAN-2000; 2000US-0483672.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
PI Kalos MD, Fanger CR, Day CH, Retter MW, Stolk JA, Skeiky YAW;  
PI Wang A, Meagher MJ;  
XX  
XX WPI; 2001-425873/45.  
XX  
XX New polynucleotide encoding a prostate-specific protein, for  
PT diagnosing, monitoring and treating prostate cancer in a patient and  
PT for use in vaccines -  
XX  
XX Claim 2; Page 464-466; 543pp; English.  
XX  
XX The present invention describes polynucleotide sequences (I) which encode  
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
CC and can be used in vaccine production and gene therapy. (I), (II),  
CC antibodies to (II), fusion proteins comprising (II), and isolated  
CC T cells prepared using (I) or (II) are used treat cancer in a patient.  
CC (I) and the antibodies are also used in the detection of cancer in a  
CC patient. The cancer that is diagnosed or treated is particularly  
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
CC (I) can be used for monitoring the progression of cancer in a patient.  
CC (I) and (II) can also be used to improve diagnostic and therapeutic  
CC methods for prostate cancer. They can indicate the level of metastasis  
CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to  
CC AAH01318 represent polynucleotide and amino acid sequences used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 371 AA;

Alignment Scores:  
Pred. No.: 3,72e-99 Length: 371  
Score: 1420.50 Matches: 275  
Percent Similarity: 88.58% Conservative: 12  
Best Local Similarity: 84.88% Mismatches: 23  
Query Match: 22.13% Indels: 14  
DB: 22 Gaps: 1

US-09-759-143-110 (1-3410) x AAH01230 (1-371)

QY 347 AACCTGTAACCTTTGGCGAGGTGCTTTGGCGAGGCATACCTATGTGCGCGCT 406  
DB 9 SerLeuValProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrSer 28  
QY 407 CTGCTGTGGGAAGTGGGGGTAGAGGAGAGTTCATGACCATGGTGTGGCGCATGGTCCA 466  
DB 29 LeuLeu-----AlaGlyIleGlyPro 35  
QY 467 GTGCTGGCGCTGCTGTGCTCCCGCTCTAGGCTCAGCCAGTACCTGGCGTGGACGC 526  
DB 36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55  
QY 527 TATGGCGCGCGCGCGCTTCTATGCGCACTGCTTGGGCATCTGCTGAGCGCTCTTT 586  
DB 56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75  
QY 587 CTCATCCCAAGGGCGGCTGCTAGCAGGGCTGTGTGCTGCCGGATCCAGGCCCTGGAG 646

DB 76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspArgProLeuGlu 95  
QY 647 CTGGCACTGCTCATCTGGGGCTGGGGCTGTGGACTTCTGTGGCCAGGTGTGCTTACT 706  
DB 96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115  
QY 707 CCACCTGGAGGCGCTGCTCTGACCTCTTCCGGGACCGGACCACTGTGCGCAGCCCTAC 766  
DB 116 ProLeuGluAlaLeuLeuSerAspLeuPheArgProAspHisCysArgGlnAlaTyr 135  
QY 767 TCTGTCTATGCTTTCATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCCATT 826  
DB 136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155  
QY 827 GACTGGGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886  
DB 156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175  
QY 887 CTGCTCACCTCATCTTCTCCTACCTGCGTAGCAGCCACACTGCTGCTGCTGCTGCTGCT 946  
DB 176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluAla 195  
QY 947 GCGCTGGGCGCCACCGAGCCAGCAGAGGGCTGTGCGGCCCTCTTGTGCGCCCACTGTC 1006  
DB 196 AlaLeuGlyProThrGluProAlaGlyLeuSerAlaProSerLeuSerProHisCys 215  
QY 1007 TGTCCATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1066  
DB 216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235  
QY 1067 CAGCTGTGCTGCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1126  
DB 236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255  
QY 1127 TGTGAGGCACTCATGACCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1186  
DB 256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275  
QY 1187 CAGGGCGTCCAGAGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 1246  
DB 276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295  
QY 1247 CGCATGGCGAGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1306  
DB 296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProGluThrThrLeuGlyAl 315  
QY 1307 ATGGACCGGC 1316  
DB 315 aValSerGly 318  
RESULT 23  
ABB95335  
ID ABB95335 standard; Protein; 371 AA.  
AC ABB95335;  
XX  
DT 19-JUL-2002 (first entry)  
XX  
DE Human P553S splice variant encoded protein SEQ ID NO 708.  
XX  
KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
KW gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN US2002022248-A1.  
XX  
XX 21-FEB-2002.  
XX  
PF 12-JAN-2001; 2001US-0759143.  
XX  
XX 25-FEB-1997; 97US-0806099.





307 gProProGluThrThrLeuGlyAlaValSerGly 318

Db

RESULT 25

ABG64105

ID ABG64105 standard; Protein; 371 AA.

XX

AC

ABG64105;

XX

DT

XX

DE

XX

KW

KW

KW

KW

KW

KW

KW

KW

OS

OS

XX

PN

XX

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

27-AUG-2002 (first entry)  
Human albumin fusion protein #780.  
Albumin fusion protein; therapeutic protein X; human albumin; HA;  
human serum albumin; HSA; cancer; reproductive disorder;  
digestive disorder; immune disorder; endocrine disorder;  
haematopoietic disorder; neural disorder; connective disorder;  
cytostatic; anti-infectivity; anti-inflammatory; anti-ulcer;  
immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;  
neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
osteopathic; antiarthritic.

Homo sapiens.

Synthetic.

WO200177137-A1.

18-OCT-2001.

12-APR-2001; 2001WO-US11988.

12-APR-2000; 2000US-229358P.

25-APR-2000; 2000US-199384P.

21-DEC-2000; 2000US-256931P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Haseltine WA;

WPI; 2002-010886/01.

New fusion protein for treating disease e.g. diabetes comprises an

albumin fused to a therapeutic protein -

Claim 1; Page 1018-1019; 2102pp; English.

The present invention relates to albumin fusion proteins comprising a  
therapeutic protein X and human albumin (HA, also known as human serum  
albumin, HSA). The proteins are useful for treating a disease or  
disorder that may be modulated by therapeutic protein X. The albumin  
extends the shelf-life of protein X, and may increase its biological  
in vitro/in vivo activity. The protein is useful for treating and  
diagnosing disorders such as cancer, reproductive disorders, digestive  
disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders  
(e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders  
(e.g. diabetes), haematopoietic disorders, neural disorders  
(e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,  
encephalomyelitis, meningitis, schizophrenia), and connective disorders  
(e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin  
fusion proteins of the invention.

Sequence 371 AA;

Alignment Scores:

Pred. No.:	4, 83e-99	Length:	371
Score:	1419.00	Matches:	269
Percent Similarity:	95.21%	Conservative:	9
Best Local Similarity:	92.12%	Mismatches:	13
Query Match:	22.11%	Indels:	1
DB:	23	Gaps:	0

US-09-759-143-110 (1-3410) x ABG64105 (1-371)

QY 443 ACCATGCTGCTGGCATTGGTCCAGTGTGGCGCTGCTGTGTCTCGGCTCTAGGCTCA 502

Db	28	SerLeuLeuAlaGlyIleGlyProValLeuGlyLeuValCysValProLeuLeuGlySer	47
QY	503	GCAGTGACCACTGGCGTGGACGCTATGGCCCGCGCGCTTCATCTTCGACCTGCTCC	562
Db	48	AlaSerAspHisTrpArgGlyArgTrpGlyArgArgProPheIleTrpAlaLeuSer	67
QY	563	TTGGGCATCCTGCTGAGCCTCTTTTCATCCCAAGGCGCGCTGCTAGCAGGGTGTCTG	622
Db	68	LeuGlyIleLeuLeuSerLeuPheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeu	87
QY	623	TGCGCGGATCCCAAGCGCGCTGGAGCTGGACCTCTCATCTGCGGCTGGGCTGCTGGAC	682
Db	88	CysProAspProArgProLeuGluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAsp	107
QY	683	TTCTGTGGCCAGGCTGTCTTCACTCCACTGGAGCGCTCTCTCTGACCTCTTCGCGGAC	742
Db	108	PheCysGlyGlnValCysPheThrProLeuGluAlaLeuSerAspLeuPheArgAsp	127
QY	743	CCGGACACTGTCCGCGGCTACTCTGTCTATGCTTCTCATGATCATGCTTGGGGCTGC	802
Db	128	ProAspHisCysArgGlnAlaTyrSerValTyrAlaPheMetIleSerLeuGlyGlyCys	147
QY	803	CTGGGCTACTCTGCTGCTCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT	862
Db	148	LeuGlyTyrLeuLeuProAlaIleAspTrpAspThrSerAlaLeuAlaProTyrLeuGly	167
QY	863	ACCCAGGAGGAGTGCCTTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT	922
Db	168	ThrGlnGluGluCysLeuPheGlyLeuLeuThrLeuIlePheLeuThrCysValAlaAla	187
QY	923	ACACTGCTGTGGCTGAGGAGGAGGCTGGGCGCCACCGAGCCAGCAGAGGCTGTCTG	982
Db	188	ThrLeuLeuValAlaGluAlaLeuLeuGlyProThrGluProAlaGluGlyLeuSer	207
QY	983	GCGCCCTCTTGTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1042
Db	208	AlaProSerLeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeu	227
QY	1043	GGCGCCCTGCTTCCCGGCTGTCACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1102
Db	228	GlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgArg	247
QY	1103	CTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1162
Db	248	LeuPheValAlaGluLeuCysSerTrpMetAlaLeuMetThrPheThrLeuPheThr	267
QY	1163	GATTTCTGCTGGCGAGGCGCTGTACCAAGGCGTGGCGGCTGAGCGGCGGCGGCGGCGG	1222
Db	268	AspPheValGlyGluGlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAla	287
QY	1223	CGGAGACACTATGATGAAGCGTTCGGATGGGAGGCTGCGGCTGCTGCTGCTGCTGCTG	1282
Db	288	ArgArgHisTyrAspGluGlyLysAlaLeu-AlaAlaSerArgGlyTrpCysGlySerAr	307
QY	1283	ATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1316
Db	307	gProProGluThrThrLeuGlyAlaValSerGly	318

RESULT 26

AAU69907

ID AAU69907 standard; Protein; 400 AA.

XX

AC

AAU69907;

XX

DT

XX

DE

XX

KW

KW

XX

OS

30-JAN-2002 (first entry)

Human prostate protein/M. tuberculosis Ra12 fusion protein RA12-P501S-E2.

Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen;  
fusion protein.

Chimeric - Homo sapiens.



PT New polynucleotide encoding a prostate-specific protein, for  
PT diagnosing, monitoring and treating prostate cancer in a patient and  
PT for use in vaccines .

PS Claim 8; Page 504-506; 543pp; English, XX

The present invention describes polynucleotide sequences (I) which encode prostate-specific proteins (II). (I) and (II) have cytostatic activity, and can be used in vaccine production and gene therapy. (I), (II), antibodies to (II), fusion proteins comprising (II), and isolated T cells prepared using (I) or (II) are used treat cancer in a patient. (I) and the antibodies are also used in the detection of cancer in a patient. The cancer that is diagnosed or treated is particularly prostate cancer. (I) and (II) can be used in vaccines. The antibodies or (I) can be used for monitoring the progression of cancer in a patient. (I) and (II) can also be used to improve diagnostic and therapeutic methods for prostate cancer. They can indicate the level of metastasis as well as the prostate volume. AAH93357 to AAH93944 and AA01115 to AA011318 represent polynucleotide and amino acid sequences used in the exemplification of the present invention.

AA	
SQ	Sequence
	400 AA;

Alignment Scores:

Fragment Size:	7.39e-98	Length:	400
Pred. No.:	1403.50	Matches:	267
Score:	97.45%	Conservative:	0
Percent Similarity:	97.45%	Mismatches:	2
Best Local Similarity:	97.45%	Indels:	5
Query Match:	21.8%	Gaps:	1
DB:	22		

US-09-759-143-110 (1-3410) x AAM01262 (1-400)

371	QY	GTGTGTTTGGCGCCGAGGC-----ATCACCTATGTGCGCGCTCTGCTGCTG	4115
127	Db	 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu	146
416	QY	GAAGTGGGGGTAGAGGAGAGATTCAACACATGCTCTGGGATTTGGTCCACAGCTGCTGGGC	475
147	Db	 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly	166
476	QY	CTGTCTCTGTGCCCGCTCTAGGCTCAGCCAGTGCACACTGGCGTGGAGCGCATGGCGCGC	535
167	Db	 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg	186
536	QY	CGCCGGCGCTTCATCTGGGCACATGCTCTTGGGCATCCTGCTGAGCGCTCTTCTCATCCCA	595
187	Db	 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro	206
596	QY	AGGCCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCGCTTGAGCTGGCACTG	655
207	Db	 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu	226
656	QY	CTCATCTCTGGCGTGGGGCTCTGGACTTCTGTGGCCAGTGTGCTTCACTCCACTGGAG	715
227	Db	 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu	246
716	QY	GCCCTGCTCTGTGACCTCTTCCGGGGACCGGACCACTGTCCGCGAGCGCTACTCTCTAT	775
247	Db	 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr	266
776	QY	GCCTTCATGATCAGCTTTGGGGGCTGCCTGGGCTACCTCTCGCTGCCATTTGACTGGAC	835
267	Db	 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp	286
836	QY	ACCAGTCCCTTGGCCCCCTACCTGGGCACCCAGGAGAGTGGCTCTTTGGGCTGCTCACC	895
287	Db	 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr	306
896	QY	CTCATCTCTCACTCGCTAGCAGCCACACTGCTGGTGGCTGAGAGCGACCGCTGGGC	955
307	Db	 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaLeuGly	326

PA (LISX/) LI S X.  
PA (WANG/) WANG A.  
PA (SKEI/) SKEIKY Y A W.  
PA (HEPL/) HEPLER W T.  
PA (HEND/) HENDERSON R A.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;  
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2002-255649/30.

XX New prostate-specific polynucleotides for diagnosing and treating  
PT diseases, in particular prostate cancer, and as markers for the  
PT progression of cancer

XX Example 17; SEQ ID NO 852; 87pp; English.

XX The present invention provides prostate-specific coding sequences and  
CC their encoded proteins. These can be used in the diagnosis and treatment  
CC of cancers, particularly prostate cancer. The present sequence is a  
CC protein described in the invention.

XX Sequence 400 AA;

#### Alignment Scores:

Pred. No.: 7.39e-98 Length: 400  
Score: 1403.50 Matches: 267  
Percent Similarity: 97.45% Conservations: 0  
Best local Similarity: 97.45% Mismatches: 2  
Query Match: 21.87% Indels: 5  
DB: 23 Gaps: 1

US-09-759-143-110 (1-3410) x ABB95367 (1-400)

QY 371 GTGCTGTTGGCGCAGGC-----ATCACCTATGTGCGCGCTCTGCTGCTG 415  
DB 127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProLeuLeuLeu 146  
QY 416 GAAGTGGGGTAGAGAGAGTTCATGACCATGGTGGTGGGCGTGGTGGTGGGCG 475  
DB 147 GluValGlyValGluGluGlyPheMetThrMetValGluGlyIleGlyProValLeuGly 166  
QY 476 CTGGCTGTGTCCCGCTCTAGGCTCAGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 535  
DB 167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTyrArgGlyArgTyrGlyArg 186  
QY 536 CGCCGGCCCTTATCTGGGCACTGCTCTGGGCGATCCTGCTGAGCCTCTTCTCATCCCA 595  
DB 187 ArgArgProPheIleThrAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206  
QY 596 AGGCGCGCTGGCTAGCAGGCTGTGTCGCCGGATCCAGCGCCCTGGAGCTGGCAGT 655  
DB 207 ArgAlaGlyTrpLeuAlaGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 226  
QY 656 CTCATCTGGCGGTGGGCTGTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715  
DB 227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246  
QY 716 GCCTGTCTGTACCTCTTCGGGACCCGGACCACTGTGGCAGGCGCTACTGTCTAT 775  
DB 247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266  
QY 776 GCCTTATGATGATCTTGGGGCTCCCTGGGCTACCTCTGCTGCCATTCACATGGGAC 835  
DB 267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286  
QY 836 ACCAGTGCCTGCCCGCTTACCTGGGACCCAGGAGGAGTGTGCTTTCGCTGTCTAC 895  
DB 287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPheGlyLeuLeuThr 306  
QY 896 CTCATCTTCTACCTGCGGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCGCGTGGGC 955  
|||||

Db 307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaLeuGly 326  
QY 956 CCACCGAGGAGCAGAGGCTGTCGCCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015  
Db 327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346  
QY 1016 CGGCGCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCGGCTGACACGCTGTC 1075  
Db 347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366  
QY 1076 TGCCGCGATGCGCCCGCAGCTGCGCGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1135  
Db 367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuLeuLeuLeuLeuLeu 386  
QY 1136 CTCATGACCTTACGCTGTTTACACGATTTTCGTCGGCGAG 1177  
Db 387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400

#### RESULT 29

AAW85068

ID AAW85068 standard; protein; 255 AA.

XX AAW85068;

XX 12-FEB-1999 (first entry)

DE Protein encoded by the consensus sequence of the PS108 gene.

XX PS108 gene; prostate disease; benign prostatic hyperplasia; BPH;

KW prostatitis; prostatic intraepithelial neoplasia; PIN; cancer;

KW drug screening; gene therapy.

XX Homo sapiens.

XX WO9805067-A1.

XX 12-NOV-1998.

XX 01-MAY-1998; 98WO-US08930.

XX 02-MAY-1997; 97US-0850713.

XX (ABBO ) ABBOTT LAB.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;

PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;

PI Roberts-Rapp L, Russell JC, Stroupe SD;

XX WPI; 1999-034731/03.

DR N-PSDB; AAW71181.

XX New isolated prostate-specific polynucleotides - used to develop

PT products for the diagnosis and treatment of prostate diseases, e.g.

PT benign hyperplasia, prostatic or prostate cancer

XX Claim 17; Pages 99-100; 122pp; English.

PS The present sequence is encoded by the consensus sequence for a PS108  
CC gene. The sequences are derived from overlapping clones AAW71166-79. The  
CC clone sequences are PS108 gene-specific. They are used in the method  
CC of the invention. The specification describes a method for detecting the  
CC presence of a target PS108 polynucleotide in a test sample. The method  
CC comprises contacting the test sample with at least 1 PS108-specific  
CC polynucleotide or complement, and detecting the presence of the target  
CC PS108 polynucleotide. The products can be used for detecting,  
CC diagnosing, staging, monitoring, prognosticating, in vivo imaging,  
CC preventing or treating, or determining predisposition to diseases or  
CC conditions of the prostate such as benign prostatic hyperplasia (BPH),  
CC prostatitis, prostatic intraepithelial neoplasia (PIN) and cancer. In  
CC particular the products can be used in drug screening and gene therapy.

XX Sequence 255 AA;

SQ



## Alignment Scores:

Pred. No.: 4,32e-89 Length: 255  
 Score: 1287.00 Matches: 255  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 20.05% Indels: 0  
 DB: 20 Gaps: 0

US-09-759-143-110 (1-3410) x AAB29268 (1-255)

QY 1178 GGGCTGTACAGCGCTGCCAGAGCTGAGCGGGCCAGCGAGCGCGGAGACACTATGAT 1237  
 DB 1 GlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgHisTyrAsp 20  
 QY 1238 GAAGCGCTTCGGATGGCGAGCTGGGGCTTCCTGCGAGTGGCCATCTCCCTGGTCTTC 1297  
 DB 21 GluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPhe 40  
 QY 1298 TCTCTGGTCTATGACCGCTGGTGCAGCATTCGGCAGCTCGAGCATCTATTGGCCAGT 1357  
 DB 41 SerLeuValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSer 60  
 QY 1358 GTGGCAGCTTTCCTGTGGCTGCCGGTGCACATGCCCTGTCCACAGATGTGGCCCTGGTG 1417  
 DB 61 ValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValVal 80  
 QY 1418 ACAGCTTCAGCGCCCTACCGGGTTCACCTTCTCAGCCCTGCAGATCTGCCCTACACA 1477  
 DB 81 ThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThr 100  
 QY 1478 CTGGCTCTCCCTTACACCGGAGAGCAGGTGTCTCTGCCCAATACCGAGGGGACACT 1537  
 DB 101 LeuAlaSerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThr 120  
 QY 1538 GGAGTGTCTAGCAGTGCAGACAGCTGTATGACAGCTTCTCCAGCGCCCTAAGCCTGGA 1597  
 DB 121 GlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGly 140  
 QY 1598 GCTCCCTTCCCTAATGGACACGTGGTGTCTGGAGCAGTGGCTCTCCACCTCCACCC 1657  
 DB 141 AlaProPheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProPro 160  
 QY 1658 GGCTCTGGGGGCCCTCTCCCTGTGATGTCTCCGTACGTGTGGTGGTGGTGGAGCCACC 1717  
 DB 161 AlaLeuCysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThr 180  
 QY 1718 GAGCCAGCGGTGTTCCGGCGGGCGGCATCTCCCTGGACCTGCCATCTGGATAGTCCC 1777  
 DB 181 GluAlaArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAla 200  
 QY 1778 TTCCTCTCTCCAGGTGGCCCATCCCTGTTTATGGCTCCATTGTCCAGCTCAGCCAG 1837  
 DB 201 PheLeuLeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGln 220  
 QY 1838 TCTGTACTGCTATATGCTGTCTCCCGAGCCCTGGGTCTGGTGGCCATTTACTTTCCT 1897  
 DB 221 SerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAla 240  
 QY 1898 ACACAGGTAGTATTTCACAGAGCCACTGGCCCAATACTAGCG 1942  
 DB 241 ThrGlnValValPheAspLysSerAspLeuAlaLysTyrSerAla 255

RESULT 30

AAB29268

ID AAB29268 standard; Protein; 255 AA.

XX AC AAB29268;

XX DT 07-FEB-2001 (first entry)

XX DE Human prostate-related PS108 partial protein sequence SEQ ID NO: 36.

XX DE Human; prostate cancer; PS108; antibody; tumour; metastasis.

XX KW

XX OS Homo sapiens.  
 XX US6130043-A.  
 XX PD 10-OCT-2000.  
 XX PF 01-MAY-1998; 98US-0071710.  
 XX PR 02-MAY-1997; 97US-0850713.  
 XX PA (ABBO ) ABBOTT LAB.  
 XX PI Cohen M, Colpitts TL, Friedman PN, Gordon J, Granados EN;  
 PI Billings-Medel PA, Klass MR, Roberts-Rapp L, Stroupe SD, Yu H;  
 PI Kratochvil JD, Russell JC, Hodges SC;  
 XX WPI: 2000-655655/63.  
 DR Methods for detecting target prostate-specific polynucleotides or  
 PT diseases of the prostate (e.g. prostate cancer), comprising detecting  
 PT the presence of any of PS108 nucleic acid sequences in a test sample -  
 XX Example 1; Column 85-88; 55pp; English.  
 CC The present invention is related to a number of partial coding and  
 CC protein sequences for the human prostate tissue protein PS108. These  
 CC sequences can be used in the diagnosis and prognosis of prostate  
 CC diseases, particularly prostate cancer. They can also be used to produce  
 CC antibodies which can be used in treatment. The present sequence is one  
 CC of the PS108 partial protein sequences.  
 XX SQ Sequence 255 AA;

## Alignment Scores:

Pred. No.: 4,32e-89 Length: 255  
 Score: 1287.00 Matches: 255  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 20.05% Indels: 0  
 DB: 21 Gaps: 0

US-09-759-143-110 (1-3410) x AAB29268 (1-255)

QY 1178 GGGCTGTACAGCGCTGCCAGAGCTGAGCGGGCCAGCGAGCGCGGAGACACTATGAT 1237  
 DB 1 GlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgHisTyrAsp 20  
 QY 1238 GAAGCGCTTCGGATGGCGAGCTGGGGCTTCCTGCGAGTGGCCATCTCCCTGGTCTTC 1297  
 DB 21 GluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPhe 40  
 QY 1298 TCTCTGTCTATGACCGCTGGTGCAGCATTCGGCAGTTCGAGCATCTATTTGGCCAGT 1357  
 DB 41 SerLeuValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSer 60  
 QY 1358 GTGGCAGCTTTCCTGTGGCTGCCGGTGCACATGCCCTGTCCACAGTGTGGCCGTGGTG 1417  
 DB 61 ValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValVal 80  
 QY 1418 ACAGCTTCAGCGCCCTACCGGGTTCACCTTCTCAGCCCTGCAGATCTCCCTACACA 1477  
 DB 81 ThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThr 100  
 QY 1478 CTGGCCTCTCTTACACCGGAGAGCAGGTGTCTCTGCCCAATACCGAGGGGACACT 1537  
 DB 101 LeuAlaSerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThr 120  
 QY 1538 GGAGTGTCTAGCAGTGCAGACAGCTTCATCAGCAGCTTCTCCAGCGCCCTAAGCCTGGA 1597  
 DB 121 GlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGly 140  
 QY 1598 GCTCCCTTCCCTAATGGACACGTGGTGTCTGGAGCAGTGGCTCTCCACCTCCACCC 1657



RESULT 32  
AAE01423  
ID AAE01423 standard; Protein: 252 AA.  
XX AC AAE01423;  
XX AC AAE01423;  
DT 18-JUL-2001 (first entry)  
XX Human secreted protein variant, SEQ ID NO:147.  
DE  
XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
KW angiotensin disorder; kidney disorder; gastrointestinal disorder; allergy;  
KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
KW cell culture; chemotaxis; vulnerability; binding partner identification;  
KW gene therapy.  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 162 /note= "Corresponds to any of the naturally occurring  
FT L-amino acids"  
FT Misc-difference 222 /note= "Corresponds to any of the naturally occurring  
FT L-amino acids"  
FT Misc-difference 236 /note= "Corresponds to any of the naturally occurring  
FT L-amino acids"  
FT Misc-difference 240 /note= "Corresponds to any of the naturally occurring  
FT L-amino acids"  
FT Misc-difference 240 /note= "Corresponds to any of the naturally occurring  
FT L-amino acids"  
WO200134629-A1.  
PD 17-MAY-2001.  
XX  
XX 08-NOV-2000; 2000WO-US30654.  
XX 12-NOV-1999; 99US-0164835.  
PR 27-JUL-2000; 2000US-0221142.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Ruben SM, Komatsoulis GA, Wei P, Fiscella M, Baker KP;  
PI WPI: 2001-308779/32.  
XX  
XX New nucleic acid encoding one of 21 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions,  
PT such as autoimmune disease and cancer, and used as a food additive or  
PT preservative -  
XX  
XX Disclosure; Page 476-477; 490pp; English.  
PS  
XX  
XX AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted  
CC protein genes, and AAE01352-AAE01413 represent the proteins they encode.  
CC AAE01415-AAE01433 represent human secreted protein fragments or variants.  
CC The secreted proteins and their genes are useful for preventing, treating  
CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
CC Pathological conditions can be diagnosed by determining the amount of the  
CC new protein in a sample or by determining the presence of mutations in  
CC the new genes. Specific uses are described for each of the 21 genes,  
CC based on the tissues in which they are most highly expressed, and include  
CC developing products for the diagnosis or treatment of proliferative  
CC disorders, cancer, tumours, foetal and developmental abnormalities,  
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),

CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
CC angiotensin disorder, kidney disorder, gastrointestinal disorders,  
CC pregnancy-related disorders, endocrine disorders, and infections. The  
CC proteins can also be used to aid wound healing and epithelial cell  
CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
CC before transplantation, for supporting cell culture of primary tissues,  
CC to regenerate tissues, to identify their cognate ligands or binding  
CC partners, and in chemotaxis, and can be used as a food additive or  
CC preservative to modify storage properties. Antibodies specific for a  
CC protein of the invention can be used in alleviating symptoms associated  
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
CC The present sequence represents a human secreted protein variant  
CC referred to in the disclosure of the invention.  
XX  
SQ Sequence 252 AA;

Alignment Scores:  
Pred. No.: 8.68e-84 Length: 252  
Score: 1217.00 Matches: 247  
Percent Similarity: 97.24% Conservative: 0  
Best Local Similarity: 97.24% Mismatches: 5  
Query Match: 18.96% Indels: 2  
DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAE01423 (1-252)

QY 1250 ATGGGACGCTGGGGCTGTTCTTCGACGTGGCCATCTCCCTGCTCTCTCTGTCATG 1309  
DB 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20  
QY 1310 GACCGCTGCTGTCAGCGAATTCGGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369  
DB 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaPhe 40  
QY 1370 CCTGTGCTGCCGGTGGCCATCGCTGTGCCACAGTGTGCCGTGTGACAGCTTCAGCC 1429  
DB 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAlaSerAla 60  
QY 1430 GCGCTCACCGGGTTCACCTTCTCAGCCCTGCACATCTGCGCTACACACTGGCTCCCTC 1489  
DB 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80  
QY 1490 TACCACCGGAGAGCAGGTGTTCTTCCCAATACCGAGGGGACACTGGAGGTGCTAGC 1549  
DB 81 TyrHisArgGlnCysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAlaSer 100  
QY 1550 AGTGAGGACAGCTGATGACACGCTTCTGCCAGGCCCTAAGCCTGGAGCTCCCTCCCT 1609  
DB 101 SerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPhePro 120  
QY 1610 AATGGACAGTGGGTGCTGGAGCAGTGGCTGCTCCACCTCCACCCCGCTCTGCGGG 1669  
DB 121 AsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeuCysGly 140  
QY 1670 GCCTGTGCTGTGATCTCTCCGTACGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1729  
DB 141 AlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAlaArgVal 160  
QY 1730 GTTCCCGGGCGGGGCTGCTGGACCTGCGCTGCTCCACCTCCACCCCGCTCTGCGGG 1789  
DB 161 Val\*\*GlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeuLeuSer 180  
QY 1790 CAGGTGGCCCCCATCTCCCTGTTTATGGGCTCCATTTGCCAGCTCAGGACGCTCTCACTGCC 1849  
DB 181 GlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerValThrAla 200  
QY 1850 TATATGCTGTGCTGCCAGGCGCTGGGTCTGGTCTGGTCTGGTCTGGTCTGGTCTGG 1909  
DB 201 TyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGlnValVal 220  
QY 1910 TTTGACAAGGCGCACTTGGCCCAAAATACTCAGCGTAGAAAACCTCCAGCACATTTGGGGTGG 1969

Db 221 Phe\*\*\*LysSerAspLeuAla-AsnThrGlnArg-AsnPheGln\*\*\*IleGlyVal\* 240  
QY 1970 AGGGCTGCTCCTACCTGGTCCAGCTCCCGCTCTCTGT 2007  
Db 240 \*\*GlyLeuProHisTrpValProAlaProCysSerCys 252

## RESULT 33

AA54369  
ID AAY54369 standard; Protein; 231 AA.

XX

AC AAY54369;

DT 06-APR-2000 (first entry)

XX Amino acid sequence of a prostate cancer-associated protein.

XX Prostate cancer-associated gene; Incyte clone 1864683; bone cancer;  
KW cell proliferation; cancer; adrenal gland cancer; bladder cancer;  
KW prostate cancer.

XX Homo sapiens.

OS

XX Key Location/Qualifiers  
FH Peptide 1..47

FT /note= "signal peptide"

FT Modified-site 100 /note= "potential casein kinase II phosphorylation site"

FT Modified-site 142 /note= "potential casein kinase II phosphorylation site"

FT Modified-site 147 /note= "potential protein kinase C phosphorylation site"

FT Domain 188..209

FT /note= "potential transmembrane domain"

XX

PN W09967384-A2.

XX

PD 29-DEC-1999.

XX

PF 15-JUN-1999; 99WO-US13524.

XX

PR 22-JUN-1998; 98US-0102615.

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Walker MG, Volkmath W, Klingler TM, Sprinzak EA;

XX

DR WPI; 2000-126631/11.

XX

DR N-PSDB; AAZ45677.

XX

PT Identifying biomolecules for the diagnosis and treatment of diseases

PT associated with cell-proliferation

XX

PS Claim 10; Page 50-51; 52pp; English.

XX

XX

CC The present sequence is encoded by a prostate cancer-associated gene,

CC whose cDNA is represented by Incyte clone number 1864683. The cDNA

CC sequence is used in the method of the invention. The specification

CC describes a method for identifying biomolecules for the diagnosis or

CC treatment of diseases associated with cell proliferation. The method

CC comprises examining polynucleotides, consisting of prostate cancer-

CC specific genes, and genes of unknown function, expressed in cDNA

CC libraries. The patterns of both gene sets are compared to identify

CC genes of unknown function with similar expression patterns to the

CC prostate cancer-specific genes. The biomolecules identified by the

CC method form pharmaceutical compositions useful for the diagnosis and

CC treatment of diseases associated with cell proliferation. Such diseases

CC include cancer of the adrenal gland, bladder and bone, but especially

CC prostate cancer. The method may also be applied using other disease-

CC specific genes. The prostate cancer-specific genes facilitate the

CC diagnosis and treatment of cell proliferation disorders.

XX

XX

XX

XX

XX

XX

XX

XX

XX

## Alignment Scores:

Pred. NO.: 1.01e-78 Length: 231  
Score: 1150.00 Matches: 230  
Percent Similarity: 99.57% Conservatives: 0  
Best Local Similarity: 99.57% Mismatches: 1  
Query Match: 17.92% Indels: 0  
DB: 21 Gaps: 0

US-09-759-143-110 (1-3410) x AAY54369 (1-231)

QY 1250 ATGGGACGCTGGGGCTGTTCTCCAGTGGCCATCTCCCTGGTCTCTCTGGTCATG 1309  
Db 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

QY 1310 GACCGGCTGGTGGAGCATTCGGCACTCGAGCAGTCTATTGGCCAGTGGGAGCTTTC 1369  
Db 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaIlePhe 40

QY 1370 CCTGTGGCTGCCGTGCCACATGCTGTCCACAGTGTGGCCGTGGTGGTGGTGGTGGT 1429  
Db 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

QY 1430 GCCCTCACCGGGTTCACTTCTCAGCCCTGCAGATCTCCCTACACACTGGCTCCCTC 1489  
Db 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

QY 1490 TACCACCGGAGAGAGAGAGTGTCTCCGCCAAATACCGAGGGGACACTGGAGGTGTAGC 1549  
Db 81 TyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAlaSer 100

QY 1550 AGTGAGGACAGCTGATGACAGCTTCTGCCAGGCCCTTAAGCCTGGAGCTCCCTTCCCT 1609  
Db 101 SerGluaspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPhePro 120

QY 1610 AATGACACAGTGGGTGGAGGAGTGGCTGCCACCTCCACCTCCACCGGCTCGCGG 1669  
Db 121 AsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeuCysGly 140

QY 1670 GCCTCTGCTGTGATCTCCGTACGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1729  
Db 141 AlaSerAlaCysAspValSerValArgValValValValValValValValValVal 160

QY 1730 GTTCCGGCGCGGGGCACTGCTGCAGCTCCCATCTCCCATCTCCCATCTCCCATCTCC 1789  
Db 161 ValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeuLeuSer 180

QY 1790 CAGGTGGCGCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGCTCTCAGTCC 1849  
Db 181 GlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerValThrAla 200

QY 1850 TATATGTTGTTCTGCCCGCAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1909  
Db 201 TyrMetValSerAlaAlaAlaLeuGlyLeuValAlaIleTyrPheAlaThrGlnValVal 220

QY 1910 TTTGACAAAGCGACTTGGCCAAATACACTACGCG 1942  
Db 221 PheAspLysSerAspLeuAlaLysTyrSerAla 231

RESULT 34  
AAY54370  
ID AAY54370 standard; Protein; 162 AA.

XX AAY54370;

XX 06-APR-2000 (first entry)

XX Amino acid sequence of a prostate cancer-associated protein.

XX Prostate cancer-associated gene; bone cancer; cell proliferation;  
KW cancer; adrenal gland cancer; bladder cancer; prostate cancer.

XX Homo sapiens.

OS

```

XX FH Key Location/Qualifiers
FT Modified-site 4
FT /note= "potential N-glycosylation site"
FT Modified-site 25
FT /note= "potential protein kinase C phosphorylation site"
FT Modified-site 34
FT /note= "potential protein kinase C phosphorylation site"
FT Modified-site 44
FT /note= "potential protein kinase C phosphorylation site"
FT Modified-site 48
FT /note= "potential cAMP and cGMP dependent protein
FT kinase phosphorylation site"
FT Domain 83..99
FT /note= "potential BLOCK signature of Ly-6/u-PAR"
XX
XX WO9967384-A2.
XX
XX 29-DEC-1999.
XX
XX 15-JUN-1999; 99WO-US13524.
XX
XX 22-JUN-1998; 98US-0102615.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Walker MG, Volkmath W, Klingler TM, Sprinzak EA;
XX
XX WPI; 2000-126631/11.
XX
XX Identifying biomolecules for the diagnosis and treatment of diseases
XX associated with cell-proliferation -
XX
XX Claim 10; Page 51-52; 52pp; English.
XX
XX The present sequence is encoded by a prostate cancer-associated gene.
XX The prostate cancer-associated gene cDNA sequence is used in the
XX method of the invention. The specification describes a method for
XX identifying biomolecules for the diagnosis or treatment of diseases
XX associated with cell proliferation. The method comprises examining
XX polynucleotides, consisting of prostate cancer-specific genes, and
XX genes of unknown function, expressed in cDNA libraries. The patterns
XX of both gene sets are compared to identify genes of unknown function
XX with similar expression patterns to the prostate cancer-specific genes.
XX The biomolecules identified by the method form pharmaceutical
XX compositions useful for the diagnosis and treatment of diseases
XX associated with cell proliferation. Such diseases include cancer of
XX the adrenal gland, bladder and bone, but especially prostate cancer.
XX The method may also be applied using other disease-specific genes. The
XX prostate cancer-specific genes facilitate the diagnosis and treatment
XX of cell proliferation disorders.
XX
XX Sequence 162 AA;
XX
XX Alignment Scores:
XX Pred. No.: 5.67e-45 Length: 162
XX Score: 704.00 Matches: 144
XX Percent Similarity: 97.35% Conservative: 3
XX Best Local Similarity: 95.36% Mismatches: 1
XX Query Match: 11.31% Indels: 3
XX DB: 21 Gaps: 0
XX
XX US-09-759-143-110 (1-3410) x AAY54370 (1-162)
XX
XX QY 447 ATGGTCATGACTTCTCTCTACCCCTCCAGCAGCAGAGCGGCACATAGGTGATG 388
XX Db 1 MetValMetAsnProSerSerThrProIleProAla-GlnArgArgHisIleGlyGln 20
XX
XX QY 387 CTGCGGCGCAACACACTCCAGCGCCAAAGTTAGCAGGTTGACACAGCAAGCTGGGCT 328
XX Db 20 er-AlaAlaLysHisThrSerArgProLysValSerArgLeuThrSerLysSerTrpAla 39
XX
XX QY 327 TTCCGGTGGCGCAGCAGCGGGCTCACCACAGGCTCTGGACCATATAGTGGGCCAGCGGGT 268

```

```

Db 40 PheArgCysArgSerArgArgLeuThrHisSerLeuTrpThrIleValGlnAlaGly 59
QY 267 AGGGCTCAGGGCGCGTTCAGGCACCTCCAGAACTGCTTCGCTCGGCTCTGCTCCAGAAG 208
Db 60 ArgAlaGlnGlyAlaValGlnAlaLeuGlnAsnCysPheValSerAlaLeuLeuGlnLys 79
QY 207 CTGCGGCTCTCTCTCTGCTGCGCCAACTGCTAGGAATCAGCCAGCGGCCCATTTCT 148
Db 80 LeuArgProLeuLeuLeuAlaAlaAsnCysLeuGlyIleSerGlnAlaProIleSer 99
QY 147 GCCAGCCCTTTGGTGGCGGCTCCAGCTTCTCAGCCCATGCTCAACACTGCTGCTGTGGGG 88
Db 100 AlaSerProLeuValProValGlnLeuLeuSerProCysSerThrProAlaAlaValGly 119
QY 87 CACCTCAGTGGGGACAGCTCTCATCACTCAGATCTGGCGGAGCGCGGCTGTCAACC 28
Db 120 HisLeuSerGlyAspThrSerHisHisSerAspProGlyArgGlyAlaArgLeuSerPro 139
QY 27 GGAGCCAGCGCGTGCAGGCTGTTCC 1
Db 140 GlyAlaSerAla***ArgLeuValPro 148
XX
XX RESULT 35
XX AAE01396
XX ID AAE01396 standard; Protein; 129 AA.
XX AC AAE01396;
XX
XX 18-JUL-2001 (first entry)
XX
XX Human gene 11 encoded secreted protein HMBAR14, SEQ ID NO:118.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
XX psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
XX inflammation; neurological disorder; Alzheimer's disease; food additive;
XX angiotensin-related disorder; kidney disorder; gastrointestinal disorder; allergy;
XX pregnancy-related disorder; endocrine disorder; infection; wound healing;
XX cell culture; chemotaxis; vulnery; binding partner identification;
XX gene therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1
XX Protein 2..129
XX /label= Signal_peptide
XX /label= Mature_human_secreted_protein
XX
XX WO200134629-A1.
XX
XX 17-MAY-2001.
XX
XX 08-NOV-2000; 2000WO-US30654.
XX
XX 12-NOV-1999; 99US-0164835.
XX
XX 27-JUL-2000; 2000US-0221142.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Wei P, Fiscella M, Baker KP;
XX WPI; 2001-308779/32.
XX
XX N-PSDB; AAD05264.
XX
XX New nucleic acid encoding one of 21 human secreted proteins for
XX diagnosing, preventing, treating or ameliorating medical conditions,
XX such as autoimmune disease and cancer, and used as a food additive or
XX preservative.
XX
XX

```

Claim 11; Page 456-457; 490pp; English.

PS AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted  
XX protein genes, and AAE01352-AAE01413 represent the proteins they encode.  
CC AAE01415-AAE01433 represent human secreted protein fragments or variants.  
CC The secreted proteins and their genes are useful for preventing, treating  
CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
CC Pathological conditions can be diagnosed by determining the amount of the  
CC new protein in a sample or by determining the presence of mutations in  
CC the new genes. Specific uses are described for each of the 21 genes,  
CC based on the tissues in which they are most highly expressed, and include  
CC developing products for the diagnosis or treatment of proliferative  
CC disorders, cancer, tumours, foetal and developmental abnormalities,  
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
CC angiotensin disorders, kidney disorders, gastrointestinal disorders,  
CC pregnancy-related disorders, endocrine disorders, and infections. The  
CC proteins can also be used to aid wound healing and epithelial cell  
CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
CC before transplantation, for supporting cell culture of primary tissues,  
CC to regenerate tissues, to identify their cognate ligands or binding  
CC partners, and in chemotaxis, and can be used as a food additive or  
CC preservative to modify storage properties. Antibodies specific for a  
CC protein of the invention can be used in alleviating symptoms associated  
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
CC The present sequence represents a human secreted protein of  
CC the invention.

XX SQ Sequence 129 AA;

#### Alignment Scores:

Pred. No.: 1.18e-42 Length: 129  
Score: 673.00 Matches: 129  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.49% Indels: 0  
DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAE01396 (1-129)

QY 2854 AGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCCACCCCTCTCTCTT 2913  
DB 1 ArgLysValGluGlyGlyAlaSerGlyLeuAsnGlyPheProAsnHisProSerSerLeu 20  
QY 2914 GGGCCAGCGCTGGTCCCGCCCACTTCCACTCCCGCTCTACTCTCTAGGACTGGGCTGATG 2973  
DB 21 GlyProAlaTrpPheProProLeuProLeuProSerThrLeuSerArgThrGlyLeuMet 40  
QY 2974 AGGGCACTGCCCAAAATTTCCCTACCCCACTTCCCTACCCCACTTCCCGCCACCC 3033  
DB 41 LysAlaLeuProLysIleSerProThrProAsnPheProLeuProProThrPheProThr 60  
QY 3034 AGCTCCACAACTCTTTGGAGCTACTGCAGGACACAGACAAAGTGCCTTTCCTCCAA 3093  
DB 61 SerSerThrThrLeuPheGlyAlaThrAlaGlyProGluAlaGlnSerAlaValSerGln 80  
QY 3094 GCCTTTGTTCATCTCAGCCCCCAGAGTATATCTGTCTGGGGAATCTCACACAAACT 3153  
DB 81 AlaPheValHisLeuSerProGlnSerIleSerValLeuGlyGluSerHisThrGluThr 100  
QY 3154 CAGGAGCACCCTCGCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGGTTAAGTGCC 3213  
DB 101 GlnGluHisProLeuProGluLeuArgGluValLeuSerLeuArgGlyGlyLeuSerAla 120  
QY 3214 GTTTCGAATAATGCTCTCTATTATT 3240  
DB 121 ValCysAsnAsnValValLeuPheIle 129

RESULT 36

ABG64102

ID ABG64102 standard; Protein; 129 AA.

XX AC ABG64102;

XX DT 27-AUG-2002 (first entry)

XX DE Human albumin fusion protein #777.

XX KW Albumin fusion protein; therapeutic protein X; human albumin; HA;  
KW human serum albumin; HSA; cancer; reproductive disorder;  
KW digestive disorder; immune disorder; endocrine disorder;  
KW haematopoietic disorder; neural disorder; connective disorder;  
KW cytostatic; antifertility; antiinflammatory; antitumor;  
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;  
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
KW osteopathic; antiarthritic.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200177137-A1.

XX PD 18-OCT-2001.

XX PF 12-APR-2001; 2001WO-US11988.

XX PR 12-APR-2000; 2000US-229358P.

XX PR 25-APR-2000; 2000US-199384P.

XX PR 21-DEC-2000; 2000US-256931P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Haseltine WA;

XX DR WPI; 2002-010886/01.

XX PT New fusion protein for treating disease e.g. diabetes comprises an  
XX albumin fused to a therapeutic protein -

XX PS Claim 1; Page 1016-1017; 2102pp; English.

XX CC The present invention relates to albumin fusion proteins comprising a  
CC therapeutic protein X and human albumin (HA, also known as human serum  
CC albumin, HSA). The proteins are useful for treating a disease or  
CC disorder that may be modulated by therapeutic protein X. The albumin  
CC extends the shelf-life of protein X, and may increase its biological  
CC in vitro/in vivo activity. The protein is useful for treating and  
CC diagnosing disorders such as cancer, reproductive disorders, digestive  
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders  
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders  
CC (e.g. diabetes), haematopoietic disorders, neural disorders  
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,  
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders  
CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin  
CC fusion proteins of the invention.

XX SQ Sequence 129 AA;

#### Alignment Scores:

Pred. No.: 1.18e-42 Length: 129  
Score: 673.00 Matches: 129  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.49% Indels: 0  
DB: 23 Gaps: 0

US-09-759-143-110 (1-3410) x ABG64102 (1-129)

QY 2854 AGGAAGGTAGAGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCCACCCCTCTCTCTT 2913

DB 1 ArgLysValGluGlyGlyAlaSerGlyLeuAsnGlyPheProAsnHisProSerSerLeu 20

QY 2914 GCGCCAGCCTGGTCCCGCCACTTCCACTCCCTCTACTCTCTAGGACTGGGCTGATG 2973  
 Db 21 GlyProAlaTrpPheProLeuProLeuProSerThrLeuSerArgThrGlyLeuMet 40  
 QY 2974 AAGGCACTGGCCCAAAATTTCCCTACCCCAACTTTCCCTACCCCAACTTTCCCAACC 3033  
 Db 41 LysAlaLeuProLysIleSerProThrProAsnPheProLeuProProThrPheProThr 60  
 QY 3034 AGCTCCACACCCCTGTTGGAGCTACTGCAGGACCAGACACAAAGTGGGTTTCCCAA 3093  
 Db 61 SerSerThrThrLeuPheGlyAlaThrAlaGlyProGluAlaGlnSerAlaValSerGln 80  
 QY 3094 GCCTTTGTCCATCTCAGCCCGCCAGAGTATATCTGTCTGGGGAATCTCACAGAACT 3153  
 Db 81 AlaPheValHisLeuSerProGlnSerIleSerValLeuGlyGluSerHisThrGluThr 100  
 QY 3154 CAGGAGCACCCCTCCCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAACTGCC 3213  
 Db 101 GlnGluHisProLeuProGluLeuArgGluValLeuSerLeuArgGlyGlyLeuSerAla 120  
 QY 3214 GTTTGCAATATGCTGCTTATTATT 3240  
 Db 121 ValCysAsnValValLeuPheIle 129

## RESULT 37

ABB60709

ID ABB60709 standard; Protein; 599 AA.

AC ABB60709;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID-NO 8919.

KW Drosophila; developmental biology; cell signalling; insecticide;

OS pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PBKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL04812.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions

PS Disclosure; SEQ ID NO 8919; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 599 AA;  
 Alignment Scores:  
 Pred. No.: 8, 24e-31 Length: 599  
 Score: 519.50 Matches: 150  
 Percent Similarity: 39.43% Conservative: 100  
 Best Local Similarity: 23.66% Mismatches: 201  
 Query Match: 8.09% Indels: 183  
 DB: 22 Gaps: 14  
 US-09-759-143-110 (1-3410) x ABB60709 (1-599)

QY 317 CGCGACCGAAGAACCCAGCTCTGTGCTCAACCTGTGCTGGAGTGTGT 376  
 Db 46 ArgLysThrArgPheGluMetPheArgLeuSerAlaIleAlaMetAlaIleGluPheAla 65  
 QY 377 TTGGCGCGAGGATCACCTATGTGCGCGCTCTGCTGCTGGAAGTGGGGTACAGGAGAG 436  
 Db 66 TyrAlaAlaGluThrSerPheValSerProIleLeuLeuGlnIleGlyValAspHisLys 85  
 QY 437 TTCATGACCATGGTGTGGCATTTGGTCCAGTGTGCTGGCGCTGTGTGCTGCCGTCTCA 496  
 Db 86 HisMetSerMetThrTrpGlyLeuSerProLeuIleGlyPhePheMetSerProLeuLeu 105  
 QY 497 GGCTCAGCCAGTACCATGGCGTGGAGCGCTATGCGCGCGCGCGCTTCTCATCTGGGCA 556  
 Db 106 GlySerIleSerAspArgCysLysLeuArgTrpGlyArgArgProIleIleSerIle 125  
 QY 557 CTGTCTTGGGCATCTGCTGAGCGCTCTTCTCATCCCA----- 595  
 Db 126 LeuSerPheGlyIleMetCysGlyLeuIleLeuValProTyrGlyLysAspLeuGlyLeu 145  
 QY 596 -----AGG 598  
 Db 146 LeuLeuGlyAspAlaGlyTyrThrTyrAlaGluSerAlaLeuAsnPheThrSerSerSer 165  
 QY 599 GCGCGCTGGCTAGCAGGCGCTGCTGTC-----CCGATCCCGAGGCC 640  
 Db 166 GlyGlySerValAlaAlaLeuValSerGlyGluAlaThrThrGlyProSerAlaSerAsp 185  
 QY 641 CTGGAGCTGGCA-----CTGCTCATCTGGCGCTGGCGCTGCTGCTGCTGGCCAG 694  
 Db 186 TyrLysPheAlaValIleLeuThrIleLeuGlyMetValLeuLeuAspPheAspAlaAsp 205  
 QY 695 GTGTCTTCATCCACTGCGAGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 754  
 Db 206 ThrCysGlnThrProAlaArgThrTyrLeuLeuAspMetCys---ValProGluGluGln 224  
 QY 755 CGCCAGGCTACTCTGCTATGCTTCTCATGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCT 814  
 Db 225 ProLysAlaMetThrMetPheAlaLeuPheAlaGlyPheGlyGlyThrIleGlyTyrAla 244  
 QY 815 CTGCTGCCATTGACTGGGACACAGTGCCTGCGCCCTACCTGCGGACCCAGGAGGAG 874  
 Db 245 IleGlyGlyValAspTrpGluThrThrHisIleGlySerPheMetGlyGlyAsnIlePro 264  
 QY 875 TGCCTCTTGGCTGCTCCTACCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928  
 Db 265 ThrValPheThrLeuValThrIleIlePheAlaValCysTyrLeuIleThrValThrThr 284  
 QY 929 -----CTGGTGGCTGAGGAGCGCGCTGGGCGCC----- 958  
 Db 285 PheArgGluIleProLeuProLeuProLeuIleGluGlnAspGluLeuLeuArgProLeuSerGlu 304  
 QY 958 ----- 958  
 Db 305 GlnAlaIleLysLysGluLeuLysLysLysAsnAsnThrIleTyrTyrIleGlnGluThr 324  
 QY 959 -----ACCGAGCCGACGAGAGG 976  
 Db 325 ThrGlnLeuGluLeuGlnMetAlaSerAspProLysArgLeuGluAlaLeuGlnGly 344  
 QY 977 CTGTGCGGCCCTCTCTTGTGCGCC----- 1000







QY	1874 GGTCTGGTCGCCATTACTTTTGCTACACAGGTAGTATTT 1912
Db	705 AlaAsnGlyValMetTyrPheSerLeuValSerPhe 717 :::     :::

RESULT 39  
AAU30352  
ID AAU30352 standard; Protein; 91 AA.

18-DEC-2001 (first entry)

Novel human secreted protein #843.

Human; vaccination; gene therapy; nutritional supplement;  
stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.  
OS  
XX  
PN WO200179449-A

AA 25-OCT-2001.

XX PF 16-APR-2001: 2001WO-IIS08656

XX  
PR 18-APR-2000. 2000US-0552929

PR 26-JAN-2001; 2001U  
yy

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 282; 765pp; English.

The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid sequences of novel human secreted proteins of the invention.

CC detecting cancer. The present sequence is a prostate specific  
 CC polypeptide of the invention.

XX  
 SQ Sequence 123 AA;

Alignment Scores:  
 Pred. No.: 6,07e-24 Length: 123  
 Score: 426.00 Matches: 87  
 Percent Similarity: 97.83% Conservative: 3  
 Best Local Similarity: 94.57% Mismatches: 2  
 Query Match: 6.64% Indels: 0  
 DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAU69873 (1-123)

QY 1250 ATGGCAGCTGGGGCTGTTCTCGAGTGGCCATCTCCCTGGTCTTCTCTGTGTCATG 1309  
 DB 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20  
 QY 1310 GACCGCTGGTGGCAGGATTCGGCACTCGAGCAGTCTATTGGCCAGTGTGGCAGCTTTC 1369  
 DB 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40  
 QY 1370 CCTGTGGCTGGCGTGCCACATGCTGCCACAGTGGCGGTGGTGCACAGCTTCAGCC 1429  
 DB 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAlaSerAla 60  
 QY 1430 GCCTCACCAGGTTACCTTCCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489  
 DB 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80  
 QY 1490 TACCACCGGAGCAGAGTGTCTCTGCCCAATAC 1525  
 DB 81 TyrHisArgGlnValLeuIleGlyGlnTrp 92

RESULT 41

ID AAM01228  
 AC AAM01228 standard; Protein; 123 AA.

XX  
 AC AAM01228;

XX  
 DT 04-OCT-2001 (first entry)

XX  
 DE P553S splice variant amino acid P553S-6.

XX  
 KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
 KW cytostatic; gene therapy; metastasis.

XX  
 OS Homo sapiens.

XX  
 PN WO200151633-A2.

XX  
 PD 19-JUL-2001.

XX  
 PF 16-JAN-2001; 2001WO-US01574.

XX  
 PR 14-JAN-2000; 2000US-0483672.

XX  
 PA (CORI-) CORIXA CORP.

XX  
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;  
 PI Wang A, Meagher MJ;

XX  
 DR WPI; 2001-425873/45.

XX  
 PT New polynucleotide encoding a prostate-specific protein, for  
 PT diagnosing, monitoring and treating prostate cancer in a patient and  
 PT for use in vaccines.

XX  
 PS Claim 2: Page 463-464; 543pp; English.

XX  
 CC The present invention describes polynucleotide sequences (I) which encode

CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
 CC and can be used in vaccine production and gene therapy. (I), (II),  
 CC antibodies to (II), fusion proteins comprising (II), and isolated  
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.  
 CC (I) and the antibodies are also used in the detection of cancer in a  
 CC patient. The cancer that is diagnosed or treated is particularly  
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
 CC (I) can be used for monitoring the progression of cancer in a patient.  
 CC (I) and (II) can also be used to improve diagnostic and therapeutic  
 CC methods for prostate cancer. They can indicate the level of metastasis  
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to  
 CC AAM01318 represent polynucleotide and amino acid sequences used in the  
 CC exemplification of the present invention.

XX Sequence 123 AA;

Alignment Scores:  
 Pred. No.: 6,07e-24 Length: 123  
 Score: 426.00 Matches: 87  
 Percent Similarity: 97.83% Conservative: 3  
 Best Local Similarity: 94.57% Mismatches: 2  
 Query Match: 6.64% Indels: 0  
 DB: 22 Gaps: 0

US-09-759-143-110 (1-3410) x AAM01228 (1-123)

QY 1250 ATGGCAGCTGGGGCTGTTCTCGAGTGGCCATCTCCCTGGTCTTCTCTGTGTCATG 1309  
 DB 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20  
 QY 1310 GACCGCTGGTGGCAGGATTCGGCACTCGAGCAGTCTATTGGCCAGTGTGGCAGCTTTC 1369  
 DB 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40  
 QY 1370 CCTGTGGCTGGCGTGCCACATGCTGCCACAGTGGCGGTGGTGCACAGCTTCAGCC 1429  
 DB 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAlaSerAla 60  
 QY 1430 GCCTCACCAGGTTACCTTCCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489  
 DB 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80  
 QY 1490 TACCACCGGAGCAGAGTGTCTCTGCCCAATAC 1525  
 DB 81 TyrHisArgGlnValLeuIleGlyGlnTrp 92

RESULT 42

ID ABB95333  
 AC ABB95333 standard; Protein; 123 AA.

XX  
 AC ABB95333;

XX  
 DT 19-JUL-2002 (first entry)

XX  
 DE Human P553S splice variant encoded protein SEQ ID NO 706.

XX  
 KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
 KW gene therapy.

XX  
 OS Homo sapiens.

XX  
 PN US2002022248-A1.

XX  
 PD 21-FEB-2002.

XX  
 PF 12-JAN-2001; 2001US-0759143.

XX  
 PR 25-FEB-1997; 97US-0806099.

XX  
 PR 01-AUG-1997; 97US-0904804.

XX  
 PR 09-FEB-1998; 98US-0020956.

XX  
 PR 25-FEB-1998; 98US-0030607.

XX  
 PR 14-JUL-1998; 98US-0115453.

XX  
 PR 23-SEP-1998; 98US-0159812.











XX PT PT

Recombinant collagen-like polymers - useful for making gels, films, fibres, etc.

XX

Example 7; Column 51; 93pp; English.

XX

This sequence represents a unnatural collagen-like polymer of the invention. The products may be used as films, fibres, moulded objects and admixed with other natural or synthetic polymers or coatings on fibres, films, labware or other surfaces, e.g. prosthetic devices. The polymers may be used for binding a wide variety of specific binding materials, as catalytic substances (where the amino acid sequence may specifically chelate a wide variety of elements), as purification media, composites, laminates or adhesives. They may also be combined with inorganic or organic materials such as carbon fibres, nylon fibres, nitrocellulose, etc., as flask coatings or in synthetic matrices for the growth and study of cells, as affinity columns or as supports for biological materials. The polymers have collagen-like properties, but may be easily expressed in micro-organisms in high efficiency. The new sequences can be tailored to give the desired properties.

XX

Sequence 762 AA;

Alignment Scores:

Pred. No.:	2,42e-21	Length:	762
Score:	395.00	Matches:	257
Percent Similarity:	34.50%	Conservative:	298
Best Local Similarity:	30.05%	Mismatches:	263
Query Match:	6.35%	Indels:	52
DB:	19	Gaps:	52

US-09-759-143-110 (1-3410) x AAW57668 (1-762)

2347	QY	GGGAACACGAGTGAGTTATTTCAGCTCCCAAAAACCCCTCTCTAGGTGTGTCT----	2291
76	Db	GlySerProGly-----AlapProGlyThrProGlyGluGlyGlnGlnHis 90	
2290	QY	CAACTAGGAGGCTAGCTGTTAACCCCTGAGCGCGGTAAATCCACCTGCA-----GAG 2240	
91	Db	HisLeuGlyGlyAlaArgGlnAlaGlyaspValGlySerProGlyAlaProGlyThrPro 110	
2239	QY	TCCCGCGAFTCCAGTCATGGAGCCCTCTCTGCCCTCCCTGTATAAGTCCAGACTGAAACC 2180	
111	Db	GlyProGlnGlyLeuProGlySerPro--GlyAlaPro-----GlyThrProGly 126	
2179	QY	CCCTGGGAAGGCTCCAGTCAGCGAGCCCTACAGACTGGGGACAGAGAGGGACGCC 2120	
127	Db	Pro--GlnGlyLeu--ProGlySerPro-----GlyAlaProGlyThrPro 140	
2119	QY	CAGCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT-----GGCAGCAGAGAG 2066	
141	Db	GlyProGlnGlyLeuPro--GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159	
2065	QY	CCAAATTACTTTGGCACACAGAAACTGTGGCGCCAGCCCGCAGCCCATCGGGCTAAAC 2006	
160	Db	GlnHisHisLeuGly-----GlyAlaArg 167	
2005	QY	AGGAGCGGGAGCTGGGACCCAGTCAGGAGCGCCCTCCACCCCAATGTCTCGAAGTTTT 1946	
168	Db	GlnAlaGlyaspValGlySerPro--GlyAlaProGlyThrPro----- 181	
1945	QY	CTAGCTGAGTATTGGCCAAGTCGCTCTGTGTAATATACTACTCTGTGTAGCAAAGTAAAT 1886	
182	Db	-----GlyProGlnGlyLeuPro----- 187	
1885	QY	GGCGACCAACCCAGGCGCTGGCGACACCAATATAGGCAGTGCACAGCTGGCTGAGCTG 1826	
188	Db	-----GlySerProGlyAlaProGlyThrProGlyProGln----- 199	
1825	QY	GACAATGGAGCCCAATAACAGGAGTGGGGCCACTGGGACAGCAGGAGGCATATCCAG 1766	
200	Db	-----GlvLeuProGlySerProGlyAla-----Prog 209	

```
Db 518 LeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro----- 533
QY 782 TGAAGGCATACACAGCTAGCCCTGGGCACAGTGT---CCGGGT---CCCGAAGAGGT 729
Db 534 -----GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 548
QY 728 CAGAGACAGCGCCCTCCAGTGGAGTGAACACACCTGGCCACAGAGTCCAGCAGGCCCA 669
Db 549 GlyThrProGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPro 566
QY 668 CGCCAGGATGAGCAGTCCAGAGTCCAGGGCCCTGGGATCCGGGCACAGCAGCCCTGCTA 609
Db 567 GlyProGlnGlyLeuProGlySerProGlyAla-----ProGlyThrProGlyGluGly 584
QY 608 GCCAGCGCGCCCTGGGATGAGAAAGA-----GGCTCAGCAGGATGCCCA 564
Db 585 GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro 604
QY 563 AGGACAGTG---CCAGATGAAGCGCGCGGCCCATAGCGTCCACGCCAGTGGTCAC 507
Db 605 GlyThrProGlyProGlnGlyLeuProGly----- 614
QY 506 TGGCTAGCCTAGAGCGGACACAGACAGCAGGCCACACTGGACCAATGCCAGCACCA 447
Db 615 -----SerProGlyAlaProGlyThrProGlyPro----- 624
QY 446 TGGTCATGAACCTCTCTACCCACCTCCAGCAGCAGAGCGCGGCACATAGGTGATGC 387
Db 625 -----GlnGlyLeuProGlySerProGlyAla----- 633
QY 386 CTGCGGCCAACACACTCCAGGCCAAAGGTAGCAGGTTGACCAAGAGCTGGGCTT 327
Db 634 -----ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr 651
QY 326 TCCGGT-----GCCGCAGCAGCGGCTCA 303
Db 652 ProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySer 671
QY 302 CCCACAGCCCTCTGACCATAGTGGCGCAGCGGG---TAGGGCTCAGGGGGCGGTCAGG 246
Db 672 ProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGly 691
QY 245 CACTCCAGAACTCTTCTCGTCTCGGCTCTGCTCCAGAAAGTGGCGGCTCTCTCTCTGCTG 186
Db 691 yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly----- 709
QY 185 CCGCCCACTCCCTAGCAATCAGCAGCGGCCCATTTCTGCCAGCCCTTTGGTSCCGGTCC 126
Db 710 -----ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGly 725
QY 125 AGCTTCTCAGCCCATCTCAACACCTGCTGCTGGGGCACCTCAGTGGGACACGCTCTC 66
Db 725 uGlyGlnGlnHisHisLeuGly-----GlyAlaArgGlnAlaGlyAspValGly 741
QY 65 ATCA-----CTCAGATCTCGGCCGA 46
Db 741 ySerProGlyAlaMet-AspProGlyArg 750
```

Search completed: June 9, 2003, 22:24:42  
Job time : 166.5 secs

Copyright (c) 1993 - 2003 CompuGen Ltd.

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3409.6	100.0	3410	19	AAV6151201	Full length cDNA s
2	3409.6	100.0	3410	19	AAV58586	Prostate tumour sp
3	3409.6	100.0	3410	21	AACT79473	cDNA sequence of h
4	3409.6	100.0	3410	22	AA5633557	Human prostate cDN
5	3409.6	100.0	3410	22	AA510108	Human prostate tum
6	3409.6	100.0	3410	22	AAH93465	Human prostate-spe
7	3409.6	100.0	3410	22	AAH84779	Human prostate-spe
8	3409.6	100.0	3410	22	AAH02530	Prostate tumour an
9	3409.6	100.0	3410	24	ABL94929	Human L1-12 cDNA s

## ALIGNMENTS

RESULT 1	
AAV61201	
ID	AAV61201 standard; cDNA; 3410 BP.
XX	
AC	AAV61201;
XX	
DT	06-JAN-1999 (first entry)
XX	
DE	Full length cDNA sequence of prostate tumour clone L1-12.
XX	
KW	Prostate; cancer; tumour; vaccine; immunogen; clone; ss.
XX	
OS	Homo sapiens.
XX	
PN	W09837093-A2.
XX	
PD	27-AUG-1998.
XX	
PF	25-FEB-1998; 98WO-US03492.
XX	
PR	09-FEB-1998; 98US-0020956.
XX	
PR	25-FEB-1997; 97US-0806099.
XX	
PR	01-AUG-1997; 97US-0904804.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Dillon DC, Xu J;
XX	
XX	WPI; 1998-609886/51.
DR	P-PSDB; AAW71869.
DR	
XX	
PT	Polypeptides comprising immunogenic portions of prostate

PT used in a vaccine for the treatment of prostate cancer

PS Claim 3; Page 79-80; 130pp; English.

xx The present sequence is a new DNA which encodes an immunogenic portion  
CC of a prostate tumour protein. The encoded immunogen, or the DNA itself,  
CC can be used as a vaccine for the treatment of prostate cancer. The DNA  
CC was identified by analysis of a subtracted cDNA library obtained by  
CC subtracting a prostate tumour cDNA expression library with a normal  
CC tissue cDNA library.

xx  
SQ sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

Query Match 100.0%; Score 3409.6; DB 19; Length 3410;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGGAACACAGCTGCACGGCTGGCTCCGGGTGACAGCGCGCCCTCGGCCAGGATCTGA	60
Db	1	GGGAACACAGCTGCACGGCTGGCTCCGGGTGACAGCGCGCCCTCGGCCAGGATCTGA	60
Qy	61	GTGATGAGAGCTGTCCCACTAGAGTGCCCAACAGACAGAGGTGTGAGCATGGGCTGAG	120
Db	61	GTGATGAGAGCTGTCCCACTAGAGTGCCCAACAGACAGAGGTGTGAGCATGGGCTGAG	120
Qy	121	AAGCTGGACCGGACCAAGGGCTGCGAGAAATGGCGCCTGGCTGATTCCTAGGAGTT	180
Db	121	AAGCTGGACCGGACCAAGGGCTGCGAGAAATGGCGCCTGGCTGATTCCTAGGAGTT	180
Qy	181	GGCGGACGACAGGAGAGGGCGGAGCTTCTGGAGCAGAGCGGAGCAAGCACTTCTG	240
Db	181	GGCGGACGACAGGAGAGGGCGGAGCTTCTGGAGCAGAGCGGAGCAAGCACTTCTG	240
Qy	241	GAGTGGCTGACGGGCCCCCTGAGCCCTACCGGCTGGCCCACTATGTTCCAGAGGCTGTG	300
Db	241	GAGTGGCTGACGGGCCCCCTGAGCCCTACCGGCTGGCCCACTATGTTCCAGAGGCTGTG	300
Qy	301	GGTGAGCGGCTGCTCGGCGACCGGAAAGCCAGCTTCTTGGTGGTCAACCTCTCAACCTT	360
Db	301	GGTGAGCGGCTGCTCGGCGACCGGAAAGCCAGCTTCTTGGTGGTCAACCTCTCAACCTT	360
Qy	361	TGGCCTGGAGGTGTGTTTGGCGCAGGCAATCACCCTATGTCGCCCTCTGCTGGAAGT	420
Db	361	TGGCCTGGAGGTGTGTTTGGCGCAGGCAATCACCCTATGTCGCCCTCTGCTGGAAGT	420
Qy	421	GGGGGTAGAGGAGAGTTCATGACACATGTGTGTGGGCAATGTCAGTGTGGGCTGTG	480
Db	421	GGGGGTAGAGGAGAGTTCATGACACATGTGTGTGGGCAATGTCAGTGTGGGCTGTG	480
Qy	481	CTGTGTCCTCCCTCCTAGGCTCAGCAGTGAACACATGGGCTGAGCCTATGGCGCGCG	540
Db	481	CTGTGTCCTCCCTCCTAGGCTCAGCAGTGAACACATGGGCTGAGCCTATGGCGCGCG	540
Qy	541	GCCCTTCATCTGGGCACTGTCTTGGGCAATCCTGCTGAGCCTCTTCTCATCCCAAGGC	600
Db	541	GCCCTTCATCTGGGCACTGTCTTGGGCAATCCTGCTGAGCCTCTTCTCATCCCAAGGC	600
Qy	601	CGGCTGGCTAGCAGGCTGTGTGCGGATCCAGGCCCCCTGGAGCTGGCACTGTCTCAT	660
Db	601	CGGCTGGCTAGCAGGCTGTGTGCGGATCCAGGCCCCCTGGAGCTGGCACTGTCTCAT	660
Qy	661	CCTGGGCTGGGCTGTGTGCTGTCAGTGTGTGTCAGTGTGTGTCAGTGTGGAGGCTT	720
Db	661	CCTGGGCTGGGCTGTGTGCTGTCAGTGTGTGTCAGTGTGTGTCAGTGTGGAGGCTT	720
Qy	721	GCTCTGTACCTTTCGGGACCCCGGACACTGTGCCAGGCTACTGTCTGTATGGCTT	780
Db	721	GCTCTGTACCTTTCGGGACCCCGGACACTGTGCCAGGCTACTGTCTGTATGGCTT	780
Qy	781	CATGATCAGTCTTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	840
Db	781	CATGATCAGTCTTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	840

```

1921 CGACTTGGCCAAATACTACGCGTAGAAATACTTCCAGCACATTTGGGTGGAGGCGCTGCCT 1980
1981 CACTGGGTCCCACTCCCGCTCCTGTAGCCGCATGGGGCTGCGGGCTGCGCCGCAGT 2040
1981 CACTGGGTCCCACTCCCGCTCCTGTAGCCGCATGGGGCTGCGGGCTGCGCCGCAGT 2040
2041 TTCTGTGTGTCGCAAAAGTAATGTGGCTCTCTGCTGCCACCCCTGCTGCTGAGGTGCGTA 2100
2041 TTCTGTGTGTCGCAAAAGTAATGTGGCTCTCTGCTGCCACCCCTGCTGCTGAGGTGCGTA 2100
2101 GCTGCACAGTGGGGCTGGGGCTGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
2101 GCTGCACAGTGGGGCTGGGGCTGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
2161 ACTGGAGGCTTCCCAAGGGGTTTCAGTCTGAGCTTATACAGGGAGGCCAAGGCTCC 2220
2161 ACTGGAGGCTTCCCAAGGGGTTTCAGTCTGAGCTTATACAGGGAGGCCAAGGCTCC 2220
2221 ATGCACCTGGAATCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
2221 ATGCACCTGGAATCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
2281 CTCCTAGTTGAGACACACCTAGAGAGGTTTTTGGGAGCTGAATAAAGCTCAGTCACCTG 2340
2281 CTCCTAGTTGAGACACACCTAGAGAGGTTTTTGGGAGCTGAATAAAGCTCAGTCACCTG 2340
2341 GTTTCCTCATCTCTAAGCCCTTAACTGCAGCTTCGTTTAAATGATAGCTTTCATGGGAG 2400
2341 GTTTCCTCATCTCTAAGCCCTTAACTGCAGCTTCGTTTAAATGATAGCTTTCATGGGAG 2400
2401 TTCTAGTAGTAACACACTCCTCCATGGGATTTGAACATATGACTTATTTAGGGGAAGA 2460
2401 TTCTAGTAGTAACACACTCCTCCATGGGATTTGAACATATGACTTATTTAGGGGAAGA 2460
2461 GTCTGAGGGGAACACACAGCAAGCAGTCCCTCAGCCACAGCAGCTGCTTTTGTCT 2520
2461 GTCTGAGGGGAACACACAGCAAGCAGTCCCTCAGCCACAGCAGCTGCTTTTGTCT 2520
2521 GATCCACCCCTCTTACCTTTTATCAGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
2521 GATCCACCCCTCTTACCTTTTATCAGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
2581 CAGAGACACAGGATTTAAATTTAACTTATTTTAACTTATTTTAACTTATTTTAACTTATTT 2640
2581 CAGAGACACAGGATTTAAATTTAACTTATTTTAACTTATTTTAACTTATTTTAACTTATTT 2640
2641 TGCTAGCTTTCTGTTGTTGCTCTAATATTTTGGGTAGGTGGGGATCCCAACAATCA 2700
2641 TGCTAGCTTTCTGTTGTTGCTCTAATATTTTGGGTAGGTGGGGATCCCAACAATCA 2700
2701 GGTCCCTGAGATAGTGTGCTATTTGGGCTGATTCATTTGCAGAAATCTTCTCTCTCTGGGT 2760
2701 GGTCCCTGAGATAGTGTGCTATTTGGGCTGATTCATTTGCAGAAATCTTCTCTCTGGGT 2760
2761 CTGGCCCCCAAAATTCCTAACCCAGGACCTTGGAAATTTACTCATCCCCAAATGATAAT 2820
2761 CTGGCCCCCAAAATTCCTAACCCAGGACCTTGGAAATTTACTCATCCCCAAATGATAAT 2820
2821 TCCAAATGCTGTATCCCAAGGTTAGGTGTTTGAAGGAAGGTAGAGGTGGGGCTTCAGGT 2880
2821 TCCAAATGCTGTATCCCAAGGTTAGGTGTTTGAAGGAAGGTAGAGGTGGGGCTTCAGGT 2880
2881 CTCACAGGCTTCCCTAACCAACCCCTCTCTCTCTGCCCCAGCTGTTCCCCCCTTCCA 2940
2881 CTCACAGGCTTCCCTAACCAACCCCTCTCTCTCTGCCCCAGCTGTTCCCCCCTTCCA 2940
2941 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3000
2941 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3000
3001 CCCAACTTCCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACC 3060

```

```

3001 CCCAACTTCCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACC 3060
3061 GCAGGACCAAGACACAAAGTGCCTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
3061 GCAGGACCAAGACACAAAGTGCCTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
3121 ATATCTGTGCTTGGGGAATCTCACACAGAACTCAGGAGCACCCTGCTGAGCTAAGG 3180
3121 ATATCTGTGCTTGGGGAATCTCACACAGAACTCAGGAGCACCCTGCTGAGCTAAGG 3180
3181 GAGGTCTTATCTCTCAGGGGGGTTAAGTGCCTTTGCAATAATGTCGCTTATTTATT 3240
3181 GAGGTCTTATCTCTCAGGGGGGTTAAGTGCCTTTGCAATAATGTCGCTTATTTATT 3240
3241 TAGCGGGGTGAATATTTTATCTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
3241 TAGCGGGGTGAATATTTTATCTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
3301 AAATTAAGGCTTCTTATATGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
3301 AAATTAAGGCTTCTTATATGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
3361 AAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3410
3361 AAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3410

```

## RESULT 2

```

AAV58586
ID AAV58586 standard; cDNA; 3410 BP.
XX AC AAV58586;
XX DT 08-DEC-1998 (first entry)
XX DE Prostate tumour specific gene clone L1-12.
XX KW Prostate tumour specific gene; human; prostate cancer; detection;
XX OS Homo sapiens.
XX FH Key
XX FT CDS
XX FT 284..1945
XX FT /*tag= a
XX PN W09837418-A2.
XX PD 27-AUG-1998.
XX PF 25-FEB-1998; 98WO-US03690.
XX PR 09-FEB-1998; 98US-0904809.
XX PR 25-FEB-1997; 97US-0806596.
XX PR 01-AUG-1997; 97US-0904809.
XX PA (CORI-) CORIXA CORP.
XX PI Dillon DC, Xu J;
XX DR WPI; 1998-480805/41.
XX DR P-ESDB; AAW69385.
XX PS Novel human prostate specific tumour protein and fragments - useful
XX PT for detecting and treating prostate cancers
XX Claim 1; Page 84-85; 14lpp; English.
XX This sequence represents a human prostate tumour specific gene, and can
XX be used in the method of the invention. The method is for detecting
XX prostate cancer comprises contacting a biological sample with an agent
XX able to bind an immunogenic portion of a prostate protein (such as
XX encoded by this sequence). An antibody which binds to an immunogenic

```

CC portion of the prostate protein, and the method can be used to detect,  
CC monitor progression of, or treat prostate cancers. The antibody may  
CC also be conjugated to a therapeutic agent for use in therapy of prostate  
XX cancers.

SQ Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

	Query Match	100.0%; Score 3409.6; DB 19; Length 3410;
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GGGAAACAGCCTGACAGCGCTGGCTCCGGGTGACAGCGCGCGCTCGGCCAGGATCTGA 60	
DB	1 GGGAAACAGCCTGACAGCGCTGGCTCCGGGTGACAGCGCGCGCTCGGCCAGGATCTGA 60	
QY	61 GTGATGACAGCTGTCACACCTGAGTGCCGCCACAGCAGCAGCTGTGAGATGGCTGAG 120	
DB	61 GTGATGACAGCTGTCACACCTGAGTGCCGCCACAGCAGCAGCTGTGAGATGGCTGAG 120	
QY	121 AAGCTGGACCGCACCAAGGCTGGCAGAAATGGCGCCTGGCTGATTCCTAGGCAAGTT 180	
DB	121 AAGCTGGACCGCACCAAGGCTGGCAGAAATGGCGCCTGGCTGATTCCTAGGCAAGTT 180	
QY	181 GCGGCACAGAGGAGAGCGCGAGCTTCTGAGCAGAGCCGAGAGCAAGCAAGTTCTG 240	
DB	181 GCGGCACAGAGGAGAGCGCGAGCTTCTGAGCAGAGCCGAGAGCAAGCAAGTTCTG 240	
QY	241 GAGTGCCTGACGGGCGCCTGAGCCCTACCGCCTGGCCACTATGGTCCAGAGGCTGTG 300	
DB	241 GAGTGCCTGACGGGCGCCTGAGCCCTACCGCCTGGCCACTATGGTCCAGAGGCTGTG 300	
QY	301 GGTGAGCGCCTGTGCGGACCGGAAAGCCAGCTTGTGTGTCAACCTGTAACTTT 360	
DB	301 GGTGAGCGCCTGTGCGGACCGGAAAGCCAGCTTGTGTGTCAACCTGTAACTTT 360	
QY	361 TGGCCTGAGGTGCTTTGGCGCAGGATCACCTATGTGCCGCTCTGCTGCTGGAAGT 420	
DB	361 TGGCCTGAGGTGCTTTGGCGCAGGATCACCTATGTGCCGCTCTGCTGCTGGAAGT 420	
QY	421 GGGGTAGAGGAGAGTTTCATGACCATGCTGGGCAATGGTCCAGTGTGGCCCTGGT 480	
DB	421 GGGGTAGAGGAGAGTTTCATGACCATGCTGGGCAATGGTCCAGTGTGGCCCTGGT 480	
QY	481 CTGTGTCGCTTCAGGCTAGGCTAGCAGTACGCTGGGCTGGAGCTATGGCCGCGCG 540	
DB	481 CTGTGTCGCTTCAGGCTAGGCTAGCAGTACGCTGGGCTGGAGCTATGGCCGCGCG 540	
QY	541 GCGCTTCATCTGGGCACTGCTCTGGGCACTCTGCTGAGCCTCTTCTCATCCCAAGGC 600	
DB	541 GCGCTTCATCTGGGCACTGCTCTGGGCACTCTGCTGAGCCTCTTCTCATCCCAAGGC 600	
QY	601 CGGCTGGCTAGCAGGCTGCTGTGCCGCGATCCCAAGCCCTGGAGCTGGCACTGCTCAT 660	
DB	601 CGGCTGGCTAGCAGGCTGCTGTGCCGCGATCCCAAGCCCTGGAGCTGGCACTGCTCAT 660	
QY	661 CTGGGCTGGGCTGCTGGACTTCTGTGGCAGGTGTCTTCACTCCAGGCGCT 720	
DB	661 CTGGGCTGGGCTGCTGGACTTCTGTGGCAGGTGTCTTCACTCCAGGCGCT 720	
QY	721 GCTCTGACCTTTCGGGACCGGACACTGTGCCAGGCTACTCTGTCTATGCTT 780	
DB	721 GCTCTGACCTTTCGGGACCGGACACTGTGCCAGGCTACTCTGTCTATGCTT 780	
QY	781 CATGATCAGTCTTGGGCTGCTGGGCTACCTCTGCTGCCATTTGACTGGGACACAG 840	
DB	781 CATGATCAGTCTTGGGCTGCTGGGCTACCTCTGCTGCCATTTGACTGGGACACAG 840	
QY	841 TGGCCTGGGCTTCTACCTGGGACCGGAGGAGTGCCTTTGGCTGCTCACTTCAT 900	
DB	841 TGGCCTGGGCTTCTACCTGGGACCGGAGGAGTGCCTTTTGGCTGCTCACTTCAT 900	
QY	901 CTTCTCAGCTGCTGAGCAGCAGCACTGCTGTGGCTGAGGAGGCGCTGGGCGCCAC 960	
DB	901 CTTCTCAGCTGCTGAGCAGCAGCACTGCTGTGGCTGAGGAGGCGCTGGGCGCCAC 960	

DB	901 CTTCTCAGCTGCTGAGCAGCAGCACTGCTGTGGCTGAGGAGGCGCTGGGCGCCAC 960
QY	961 CGAGCAGCAGAGAGGCTGTGCGGCCCTCTCTGTGTCGCCCTGCTGTCTCCATGCGGCG 1020
DB	961 CGAGCAGCAGAGAGGCTGTGCGGCCCTCTCTGTGTCGCCCTGCTGTCTCCATGCGGCG 1020
QY	1021 CCGCTTGGCTTTCGGGAACCTGGGCGCCTGCTTCCCGGCTGACACAGCTGTGTCGCG 1080
DB	1021 CCGCTTGGCTTTCGGGAACCTGGGCGCCTGCTTCCCGGCTGACACAGCTGTGTCGCG 1080
QY	1081 CATGCCCGCACCTCGCGGCTCTGCTGGCTGAGCTGTGACAGTGTGACAGTGTGACACTCAT 1140
DB	1081 CATGCCCGCACCTCGCGGCTCTGCTGGCTGAGCTGTGACAGTGTGACAGTGTGACACTCAT 1140
QY	1141 GACCTTCAGCTCTTTTACACGATTTCTGTTGGCGAGGGCTGTACAGGGGCTGCCAG 1200
DB	1141 GACCTTCAGCTCTTTTACACGATTTCTGTTGGCGAGGGCTGTACAGGGGCTGCCAG 1200
QY	1201 AGCTGAGCGGGCACCGAGCGCGGAGACACTATGATGAAGGCTTCGGATGGGACGCT 1260
DB	1201 AGCTGAGCGGGCACCGAGCGCGGAGACACTATGATGAAGGCTTCGGATGGGACGCT 1260
QY	1261 GGGCTGTTCTCAGTGGGCACTCTCCCTGCTTCTCTGCTCATGACCGGCTGT 1320
DB	1261 GGGCTGTTCTCAGTGGGCACTCTCCCTGCTTCTCTGCTCATGACCGGCTGT 1320
QY	1321 GCAGGATTCGGCACTCGAGCACTATTTGGCCAGTGTGGCAGCTTCCCTGTGGCTGC 1380
DB	1321 GCAGGATTCGGCACTCGAGCACTATTTGGCCAGTGTGGCAGCTTCCCTGTGGCTGC 1380
QY	1381 CGGTGCCACATGCTTCCACAGTGTGGCGTGGTGACAGCTTACGCGCCTCACCG 1440
DB	1381 CGGTGCCACATGCTTCCACAGTGTGGCGTGGTGACAGCTTACGCGCCTCACCG 1440
QY	1441 GTTCACCTTCTCAGCCTGCAGATCTGCGCTACACACTGGCCTCCCTTACCACCGGA 1500
DB	1441 GTTCACCTTCTCAGCCTGCAGATCTGCGCTACACACTGGCCTCCCTTACCACCGGA 1500
QY	1501 GAAGAGGCTTCTCGCCAAATACGAGGGGACACTGGAGTGTGACAGTGTGAGGACAG 1560
DB	1501 GAAGAGGCTTCTCGCCAAATACGAGGGGACACTGGAGTGTGACAGTGTGAGGACAG 1560
QY	1561 CTTGATGACAGCTTCTGCGCAGCCTTAAGCTTGGAGCTTCCCTTCCCTAATGACACGT 1620
DB	1561 CTTGATGACAGCTTCTGCGCAGCCTTAAGCTTGGAGCTTCCCTTCCCTAATGACACGT 1620
QY	1621 GGGTGTGAGGAGTGGCTTCTCCACCTCCACCGCGCTCTGCGGGGCTTCTGCCTG 1680
DB	1621 GGGTGTGAGGAGTGGCTTCTCCACCTCCACCGCGCTCTGCGGGGCTTCTGCCTG 1680
QY	1681 TGATGCTTCCGTAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740
DB	1681 TGATGCTTCCGTAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740
QY	1741 GGGATGCTGCTGACCTCGCCATCTGATGCTTCTGCTTCCAGGCTGGTGGTGGTGG 1800
DB	1741 GGGATGCTGCTGACCTCGCCATCTGATGCTTCTGCTTCCAGGCTGGTGGTGGTGG 1800
QY	1801 ATCCCTTTTATGGCTTCCATTTGTCAGCTTACGCTGTGCTTCCCTTATGCTGTC 1860
DB	1801 ATCCCTTTTATGGCTTCCATTTGTCAGCTTACGCTGTGCTTCCCTTATGCTGTC 1860
QY	1861 TGGCGAGGCTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920
DB	1861 TGGCGAGGCTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920
QY	1921 CGACTTGGCAAAATACCTACGCTAGAAACTTCCAGCACTTGGGTGGAGGCTGGCT 1980
DB	1921 CGACTTGGCAAAATACCTACGCTAGAAACTTCCAGCACTTGGGTGGAGGCTGGCT 1980
QY	1981 CACTTGGTCCAGCTCCCGCTCTCTGTAGCCCATGGGCTGCCGGCTGGCGGCTG 2040
DB	1981 CACTTGGTCCAGCTCCCGCTCTCTGTAGCCCATGGGCTGCCGGCTGGCGGCTG 2040





QY 1 GGGACACAGCTGCACGGCTGGCTCGGGTGACAGCGCGCGCTCGGCCAGATCTGA 60  
DB 1 GGGACACAGCTGCACGGCTGGCTCGGGTGACAGCGCGCGCTCGGCCAGATCTGA 60  
QY 61 GTGATGACAGCTGTCGCCACTGAGTGGCCACACAGCAGCAGTGTGTGAGCATGGCTGAG 120  
DB 61 GTGATGACAGCTGTCGCCACTGAGTGGCCACACAGCAGCAGTGTGTGAGCATGGCTGAG 120  
QY 121 AAGCTGAGCCGACCAAGGGCTGGCAGAAATGGGGCTTGGCTGATTTCTAGGCAGTT 180  
DB 121 AAGCTGAGCCGACCAAGGGCTGGCAGAAATGGGGCTTGGCTGATTTCTAGGCAGTT 180  
QY 181 GCGGCACAGAGGAGAGCGCGAGCTTCTGAGCAGAGCCAGAGCAGTGTCTG 240  
DB 181 GCGGCACAGAGGAGAGCGCGAGCTTCTGAGCAGAGCCAGAGCAGTGTCTG 240  
QY 241 GAGTCCCTGAAAGGGCCCTGAGCCCTACCGCTGCGCCACTATGTTCAGAGGCTGTG 300  
DB 241 GAGTCCCTGAAAGGGCCCTGAGCCCTACCGCTGCGCCACTATGTTCAGAGGCTGTG 300  
QY 301 GGTGAGCCGCTGTGCGGCACCGAAAGCCAGCTTGTGTGTCACACCTGTAACTT 360  
DB 301 GGTGAGCCGCTGTGCGGCACCGAAAGCCAGCTTGTGTGTCACACCTGTAACTT 360  
QY 361 TGGCTGAGGTGTGTTGGCGCAGGATCACCTATGTGCGGCTCTGTGCTGGAAGT 420  
DB 361 TGGCTGAGGTGTGTTGGCGCAGGATCACCTATGTGCGGCTCTGTGCTGGAAGT 420  
QY 421 GGGGTAGAGGAGTTCATGACCATGGTGTGGCATTTGGTCCAGTGTGGGCTGGT 480  
DB 421 GGGGTAGAGGAGTTCATGACCATGGTGTGGCATTTGGTCCAGTGTGGGCTGGT 480  
QY 481 CTGTCTCCGCTCTAGGCTGAGCAGTACCATGGTGGAGCTATGGCGCCGCGCG 540  
DB 481 CTGTCTCCGCTCTAGGCTGAGCAGTACCATGGTGGAGCTATGGCGCCGCGCG 540  
QY 541 GCGCTTCATCTGGGCACCTGCTTGGGCATCTGCTGAGCCTTCTTCATCCCAAGGC 600  
DB 541 GCGCTTCATCTGGGCACCTGCTTGGGCATCTGCTGAGCCTTCTTCATCCCAAGGC 600  
QY 601 CGGCTGGCTAGCAGGGCTGTGTCGCCGATCCAGGCCCTGGAGCTGGCAGTCTCAT 660  
DB 601 CGGCTGGCTAGCAGGGCTGTGTCGCCGATCCAGGCCCTGGAGCTGGCAGTCTCAT 660  
QY 661 CCGTGGCTGGGCTGTGGACTTCTGTGGCAGTGTGCTTCACTGAGGAGGCT 720  
DB 661 CCGTGGCTGGGCTGTGGACTTCTGTGGCAGTGTGCTTCACTGAGGAGGCT 720  
QY 721 GCTCTGACCTTTCGGGACCGGACCACTGTCGCCAGGCTTCTGTCTATGCTT 780  
DB 721 GCTCTGACCTTTCGGGACCGGACCACTGTCGCCAGGCTTCTGTCTATGCTT 780  
QY 781 CATGATCAGTCTGGGGCTGCCTGGGTACCTCTGCTGCCATTTGACTGGGACACAG 840  
DB 781 CATGATCAGTCTGGGGCTGCCTGGGTACCTCTGCTGCCATTTGACTGGGACACAG 840  
QY 841 TGGCTGGCCCTTACCTGGGACCGGAGGAGTGTGCTTGGCTGCTGCTGCTCAT 900  
DB 841 TGGCTGGCCCTTACCTGGGACCGGAGGAGTGTGCTTGGCTGCTGCTGCTCAT 900  
QY 901 CTTCTCAGCTTGGTGGGACCACTGCTGGTGGCTGAGGAGGAGGCTGGGCGCCAC 960  
DB 901 CTTCTCAGCTTGGTGGGACCACTGCTGGTGGCTGAGGAGGAGGCTGGGCGCCAC 960  
QY 961 CGAGCCAGAGAGGGCTGTGCGGCCCTCTCTGTGCGCCACTGTCTGCTGCGGGC 1020  
DB 961 CGAGCCAGAGAGGGCTGTGCGGCCCTCTCTGTGCGCCACTGTCTGCTGCGGGC 1020  
QY 1021 CGGCTTGGCTTTCGGAACCTGGGCGGCTGCTTCCCGGCTGACAGGCTGCTGCG 1080  
DB 1021 CGGCTTGGCTTTCGGAACCTGGGCGGCTGCTTCCCGGCTGACAGGCTGCTGCG 1080  
QY 1081 CATGCCCGGACCTTGGCGGCTTCTGCTGGCTGAGTGTGACAGTGGTGGACACTCAT 1140

DB 1081 CATGCCCGGACCTTGGCGGCTTCTGCTGGCTGAGTGTGACAGTGGTGGACTCAT 1140  
QY 1141 GACCTTTCAGCTCTTTTACACGATTTCTGTTGGGAGGGCTGTACAGGGCTGCCCCAG 1200  
DB 1141 GACCTTTCAGCTCTTTTACACGATTTCTGTTGGGAGGGCTGTACAGGGCTGCCCCAG 1200  
QY 1201 AGCTGAGCCGGGACCGAGCGCGGAGACACTATATGATGAAGCGTTTCGATGGCAGCT 1260  
DB 1201 AGCTGAGCCGGGACCGAGCGCGGAGACACTATATGATGAAGCGTTTCGATGGCAGCT 1260  
QY 1261 GGGCTGTTCTGCTGAGTGGCCATCTCTGCTGCTTCTCTGCTGCTATGAGCGGCTGT 1320  
DB 1261 GGGCTGTTCTGCTGAGTGGCCATCTCTGCTGCTTCTCTGCTGCTATGAGCGGCTGT 1320  
QY 1321 GCAGCATTTCGGCAGCTGAGCAGTCTATTTGGCAGTGTGGCAGTCTTCCCTGTGGCTGC 1380  
DB 1321 GCAGCATTTCGGCAGCTGAGCAGTCTATTTGGCAGTGTGGCAGTCTTCCCTGTGGCTGC 1380  
QY 1381 CGGTGCCACATGCTTGTCCACAGTGTGGCGTGTGACAGTTCAGCGCGCTTACCAG 1440  
DB 1381 CGGTGCCACATGCTTGTCCACAGTGTGGCGTGTGACAGTTCAGCGCGCTTACCAG 1440  
QY 1441 GTTCACTTCTCAGCCCTGCAGATCTGCTTACACACTGGCCTCTCTACACCGGGA 1500  
DB 1441 GTTCACTTCTCAGCCCTGCAGATCTGCTTACACACTGGCCTCTCTACACCGGGA 1500  
QY 1501 GAAGCAGGTGTTCTCTCCCAATACCGAGGGGACTGGAGTGTGTAGCAGTGGACAG 1560  
DB 1501 GAAGCAGGTGTTCTCTCCCAATACCGAGGGGACTGGAGTGTGTAGCAGTGGACAG 1560  
QY 1561 CTTGATGACAGCTTCTTCCGAGGCTTAAAGCTTGGAGTCTTCCCTTAATGGACAGT 1620  
DB 1561 CTTGATGACAGCTTCTTCCGAGGCTTAAAGCTTGGAGTCTTCCCTTAATGGACAGT 1620  
QY 1621 GGTGTGTGAGGAGTGGCTGCTCCACCTCCACCGGCTGTGGGGGCTTCCCTG 1680  
DB 1621 GGTGTGTGAGGAGTGGCTGCTCCACCTCCACCGGCTGTGGGGGCTTCCCTG 1680  
QY 1681 TGATGTCGCTGACGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
DB 1681 TGATGTCGCTGACGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
QY 1741 GGGCATCTGCTGAGCTTCCCATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
DB 1741 GGGCATCTGCTGAGCTTCCCATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
QY 1801 ATCCCTGTTTATGGGCTTCCATTTGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTG 1860  
DB 1801 ATCCCTGTTTATGGGCTTCCATTTGCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTG 1860  
QY 1861 TGCCGAGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
DB 1861 TGCCGAGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
QY 1921 CGACTTGGCCAAATCTCAGCTGAGGAGTTCACACACTTGGGCTGAGGAGGCTGCTGCT 1980  
DB 1921 CGACTTGGCCAAATCTCAGCTGAGGAGTTCACACACTTGGGCTGAGGAGGCTGCTGCT 1980  
QY 1981 CACTGGTCCAGCTCCCGCTCTGTTAGCCCATGGGCTGCGGGCTGCGCGGCTGCG 2040  
DB 1981 CACTGGTCCAGCTCCCGCTCTGTTAGCCCATGGGCTGCGGGCTGCGCGGCTGCG 2040  
QY 2041 TTCTGTTGCTGCAAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
DB 2041 TTCTGTTGCTGCAAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
QY 2101 GCTGACAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
DB 2101 GCTGACAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
QY 2161 ACTGGAGGCTTCCAAAGGGGTTTACGCTGAGCTTATACAGGAGGCGGAGGAGGCTCC 2220  
DB 2161 ACTGGAGGCTTCCAAAGGGGTTTACGCTGAGCTTATACAGGAGGCGGAGGAGGCTCC 2220

2161	ACTGAGGCCTTCCAAAGGGGGTTTCAGTCTCGACTTATACAGGAGGCCCAAGAGGGCTCC	2220
Db		
2221	ATGCACCTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTTAACAGCTAGC	2280
Qy		
2221	ATGCACCTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTTAACAGCTAGC	2280
Db		
2281	CTCCTAGTTGAGACACACTAGAGNAGGGTTTTTTGGGAGCTGAAATAAATCAGTCACCTG	2340
Qy		
2281	CTCCTAGTTGAGACACACTAGAGNAGGGTTTTTTGGGAGCTGAAATAAATCAGTCACCTG	2340
Db		
2341	GTTTCCCATCTCTAAGCCCCCTTAACCTGCAGCTTCGTTTTAATGTAGCTCTTTGCATGGAG	2400
Qy		
2341	GTTTCCCATCTCTAAGCCCCCTTAACCTGCAGCTTCGTTTTAATGTAGCTCTTTGCATGGAG	2400
Db		
2401	TTTCTAGGATGAAACACTCTCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA	2460
Qy		
2401	TTTCTAGGATGAAACACTCTCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA	2460
Db		
2461	GTCTCGAGGGGCAACACAGAACCAGCTCCCTCAGCCACACGACTGTCTTTTGGT	2520
Qy		
2461	GTCTCGAGGGGCAACACAGAACCAGCTCCCTCAGCCACACGACTGTCTTTTGGT	2520
Db		
2521	GATCCACCCCTCTTTACTCTTTATCAGATGTGGCTGTTGGTCTCTCTGTGTGCCATCA	2580
Qy		
2521	GATCCACCCCTCTTTACTCTTTATCAGATGTGGCTGTTGGTCTCTCTGTGTGCCATCA	2580
Db		
2581	CAGACACACAGCACTTAAATATTAAACTTATTTATTTAAACAAAGTAGAAGGGAATCCAT	2640
Qy		
2581	CAGACACACAGCACTTAAATATTAACTTATTTATTTAAACAAAGTAGAAGGGAATCCAT	2640
Db		
2641	TGCTAGCTTTTCTGTGTGGTCTCTAATAATTTGGGTAGGTTGGGGATCCCCACAATCA	2700
Qy		
2641	TGCTAGCTTTTCTGTGTGGTCTCTAATAATTTGGGTAGGTTGGGGATCCCCACAATCA	2700
Db		
2701	GGTCCCCCAGATAGCTGTCATTTGGGCTGATCATTTGCCAGAACTCTTCTCTCCGGGGT	2760
Qy		
2701	GGTCCCCCAGATAGCTGTCATTTGGGCTGATCATTTGCCAGAACTCTTCTCTCCGGGGT	2760
Db		
2761	CTGGCCCCCAAAATGCCTTAACCCAGGACCTTGGAAATTTCTACTCATCCCAATGATAAT	2820
Qy		
2761	CTGGCCCCCAAAATGCCTTAACCCAGGACCTTGGAAATTTCTACTCATCCCAATGATAAT	2820
Db		
2821	TCCAAATGCTGTTTACCCAGGTTAGGGTGTGAGGAAGGTAGAGGGTGGGGCTTCAGGT	2880
Qy		
2821	TCCAAATGCTGTTTACCCAGGTTAGGGTGTGAGGAAGGTAGAGGGTGGGGCTTCAGGT	2880
Db		
2881	CTCAACGGCTTCCTTAACACACCCCTCTCTCTTGGCCAGCCCTTGTTGCCCCACATTCCA	2940
Qy		
2881	CTCAACGGCTTCCTTAACACACCCCTCTCTCTTGGCCAGCCCTTGTTGCCCCACATTCCA	2940
Db		
2941	CTCCCCCTACTCTCTCTAGGACTGGGCTGATGAAGGCATGCCCCAAAATTTCCCTTACC	3000
Qy		
2941	CTCCCCCTACTCTCTCTAGGACTGGGCTGATGAAGGCATGCCCCAAAATTTCCCTTACC	3000
Db		
3001	CCCAACTTTCCCTTACCCCAACTTTCCCAACAGCTCCACAACCCCTGTTTGGAGCTACT	3060
Qy		
3001	CCCAACTTTCCCTTACCCCAACTTTCCCAACAGCTCCACAACCCCTGTTTGGAGCTACT	3060
Db		
3061	GCAGGACCAAGACAAAGTGGGGTTTCCCAAGCCCTTTGTCCATCTCAGCCCCCAGAGT	3120
Qy		
3061	GCAGGACCAAGACAAAGTGGGGTTTCCCAAGCCCTTTGTCCATCTCAGCCCCCAGAGT	3120
Db		
3121	ATATCTGTGTTGGGAAATCTCACAGAAACTCAGGAGCACCCCTGCCTGAGCTAAGG	3180
Qy		
3121	ATATCTGTGTTGGGAAATCTCACAGAAACTCAGGAGCACCCCTGCCTGAGCTAAGG	3180
Db		
3181	GAGTCTTATCTCTCAGGGGGGTTTAAAGTGCCGTTTGTCAATAATGTCGCTTATTTATT	3240
Qy		
3181	GAGTCTTATCTCTCAGGGGGGTTTAAAGTGCCGTTTGTCAATAATGTCGCTTATTTATT	3240
Db		
3241	TAGCGGGGTGAAATATTTTATCTGTAAGTGAGCAATCAGAGTATAATGTTTTATGGTGACA	3300
Qy		
3241	TAGCGGGGTGAAATATTTTATCTGTAAGTGAGCAATCAGAGTATAATGTTTTATGGTGACA	3300
Db		

QY	3301	AAATTAAAGGCTTCTATATGTTTAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3360
Db	3301	AAATTAAAGGCTTCTATATGTTTAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3360
QY	3361	AAAAAARAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3410
Db	3361	AAAAAARAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3410
 RESULT 4 AAS63557				
ID	AAS63557	standard; cDNA; 3410 BP.		
XX	AC	AAS63557;		
XX	DT	29-JAN-2002 (first entry)		
XX	DE	Human prostate cDNA sequence #109.		
XX	KW	Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.		
XX	OS	Homo sapiens.		
XX	PX	WO200173032-A2.		
XX	PD	04-OCT-2001.		
XX	PF	27-MAR-2001; 2001WO-US09919.		
XX	PR	27-MAR-2000; 2000US-0536857.		
XX	PR	09-MAY-2000; 2000US-0568100.		
XX	PR	12-MAY-2000; 2000US-0570737.		
XX	PR	13-JUN-2000; 2000US-0593793.		
XX	PR	27-JUN-2000; 2000US-0605783.		
XX	PR	10-AUG-2000; 2000US-0636215.		
XX	PR	29-AUG-2000; 2000US-0651236.		
XX	PR	06-SEP-2000; 2000US-0657279.		
XX	PR	02-OCT-2000; 2000US-0679426.		
XX	PR	10-OCT-2000; 2000US-0685166.		
XX	PA	(CORI-) CORIXA CORP.		
XX	PI	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;		
XX	PI	Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;		
XX	PI	Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;		
XX	PS	WPI; 2001-639232/73.		
XX	DR	P-PSDB; AAU69763.		
XX	PT	New human prostate-specific polypeptides and polynucleotides useful for		
XX	PT	the diagnosis and treatment of cancer, especially prostate cancer -		
XX	PS	Claim 1; Page 267-268; 579pp; English.		
XX	CC	The invention relates to isolated prostate-specific		
XX	CC	polynucleotides, polypeptides, fusion proteins of the polypeptides,		
XX	CC	antibodies raised against the polypeptides (or antigenic epitopes		
XX	CC	derived from them) and antigen-presenting cells expressing the		
XX	CC	polypeptides. The antibodies are useful for detecting the presence of		
XX	CC	cancer, especially prostate cancer. The polypeptides, polynucleotides and		
XX	CC	the antigen-presenting cells are useful for stimulating and/or expanding		
XX	CC	T cells specific for a tumour protein, and for inhibiting the development		
XX	CC	of cancer especially prostate cancer. Compositions comprising the		
XX	CC	polynucleotide and/or polypeptide are useful for stimulating an immune		
XX	CC	response, and for treating cancer. The oligonucleotide is useful for		
XX	CC	detecting cancer. The present sequence is a prostate specific		
XX	CC	polynucleotide of the invention.		
XX	Sequence	3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;		
XX	Query Match	100.0%; Score 3409.6; DB 22; Length 3410;		
XX	Best Local Similarity	100.0%; Pred. No. 0;		

Query Match 100.0%; Score 3409.6; DB 22; Length 3410;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 3410;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------------	--------------	----	------------	----	--------	----	------	----

[illegible]

QY 2161 ACTGAGGCTTCAAGGGGTTTCAGTCTGAGCTTATACAGGAGGCCAGAGGGCTCC 2220  
 Db 2161 ACTGAGGCTTCAAGGGGTTTCAGTCTGAGCTTATACAGGAGGCCAGAGGGCTCC 2220  
 QY 2221 ATGCACTGGAATGGGGGACTCTGAGGTGGATTTACCCAGGCTCAGGGTTAAACAGCTAGC 2280  
 Db 2221 ATGCACTGGAATGGGGGACTCTGAGGTGGATTTACCCAGGCTCAGGGTTAAACAGCTAGC 2280  
 QY 2281 CTCCTAGTTGAGACACACACCTAGAGAGGGTTTTGGGAGCTGAATAACTCAGTCACCTG 2340  
 Db 2281 CTCCTAGTTGAGACACACCTAGAGAGGGTTTTGGGAGCTGAATAACTCAGTCACCTG 2340  
 QY 2341 GTTCCCATCTCTAAGCCCTTAACCTGCAGCTTCGTTAATGTAGCTCTTGCAATGGAG 2400  
 Db 2341 GTTCCCATCTCTAAGCCCTTAACCTGCAGCTTCGTTAATGTAGCTCTTGCAATGGAG 2400  
 QY 2401 TTTCTAGGATGAACACACTCTCCATGGGATTTGAACATATGACTTATTTGAGGGGAAGA 2460  
 Db 2401 TTTCTAGGATGAACACACTCTCCATGGGATTTGAACATATGACTTATTTGAGGGGAAGA 2460  
 QY 2461 GTCTGAGGGGGAACACACAGAACAGGTCCTCCCTCAGCCACAGCAGCTGCTTTTGGCT 2520  
 Db 2461 GTCTGAGGGGGAACACACAGAACAGGTCCTCCCTCAGCCACAGCAGCTGCTTTTGGCT 2520  
 QY 2521 GATCCACCCCTCTTACCTTTATCAGGATGTGGCTGTGGCTCTCTCTGTCGCAATCA 2580  
 Db 2521 GATCCACCCCTCTTACCTTTATCAGGATGTGGCTGTGGCTCTCTCTGTCGCAATCA 2580  
 QY 2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTAATAAAGTAGAGGGGAATCCAT 2640  
 Db 2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTAATAAAGTAGAGGGGAATCCAT 2640  
 QY 2641 TCTAGCTTTCTGTGTTGGTCTTAATATTTGGTAGGGTGGGGATCCCCCAACAATCA 2700  
 Db 2641 TCTAGCTTTCTGTGTTGGTCTTAATATTTGGTAGGGTGGGGATCCCCCAACAATCA 2700  
 QY 2701 GGTCCCTCTGAGATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 2760  
 Db 2701 GGTCCCTCTGAGATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 2760  
 QY 2761 CTGGCCCTCTGAGATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 2820  
 Db 2761 CTGGCCCTCTGAGATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 2820  
 QY 2821 TCAATGCTGTTTACCAAGGTTAGGTTGTTGAAGGAAGGTAGAGGGTGGGCTCAGGT 2880  
 Db 2821 TCAATGCTGTTTACCAAGGTTAGGTTGTTGAAGGAAGGTAGAGGGTGGGCTCAGGT 2880  
 QY 2881 CTCAAGGGCTTCCCTAACACACCCCTCTCTCTTGGCCAGCCTGCTGCTGCTGCTGCTGCTG 2940  
 Db 2881 CTCAAGGGCTTCCCTAACACACCCCTCTCTCTTGGCCAGCCTGCTGCTGCTGCTGCTGCTG 2940  
 QY 2941 CTCCCTCTACTCTCTAGGACTGGGCTGATGAGGCACTGCCCAAAATTTCCCTTACC 3000  
 Db 2941 CTCCCTCTACTCTCTAGGACTGGGCTGATGAGGCACTGCCCAAAATTTCCCTTACC 3000  
 QY 3001 CCCAAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 3060  
 Db 3001 CCCAAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 3060  
 QY 3061 GCAGGACCAAGAACCAAGTGGGTTTCCCAAGCCTTTGTCATCTCAGCCCTCAGAGT 3120  
 Db 3061 GCAGGACCAAGAACCAAGTGGGTTTCCCAAGCCTTTGTCATCTCAGCCCTCAGAGT 3120  
 QY 3121 ATATCTGTGTTGGGGAATCTCACAGAAACTCAGGAGCAGCCCTGCTGAGCTAAGG 3180  
 Db 3121 ATATCTGTGTTGGGGAATCTCACAGAAACTCAGGAGCAGCCCTGCTGAGCTAAGG 3180  
 QY 3181 GAGGTCTTATCTCAGGGGGGTTTAAAGTGGCGTTTGAATAATGCTGCTTATTTAT 3240  
 Db 3181 GAGGTCTTATCTCAGGGGGGTTTAAAGTGGCGTTTGAATAATGCTGCTTATTTAT 3240  
 QY 3241 TAGCGGGTGAATATTTATATCTGTAAGTGAGCAATCAGATATAATGTTATGTTGAC 3300

Db 3241 TAGCGGGTGAATATTTATATCTGTAAGTGAGCAATCAGATATAATGTTATGTTGAC 3300  
 QY 3301 AAATTTAAAGGCTTCTTATATGTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3360  
 Db 3301 AAATTTAAAGGCTTCTTATATGTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3360  
 QY 3361 AAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAA 3410  
 Db 3361 AAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAA 3410

RESULT 5  
 AAS10108  
 ID AAS10108 standard; cDNA; 3410 BP.  
 XX AAS10108;  
 AC AAS10108;  
 XX AAS10108;  
 DT 24-OCT-2001 (first entry)  
 XX Human prostate tumour cDNA LI-12.  
 DE Human prostate tumour protein; prostate cancer; ss.  
 KW Homo sapiens.  
 XX OS  
 XX US6262245-B1.  
 XX 17-JUL-2001.  
 XX 25-FEB-1998; 98US-0030607.  
 XX 25-FEB-1997; 97US-0806099.  
 XX 01-AUG-1997; 97US-0904804.  
 XX 09-FEB-1998; 98US-0020956.  
 XX (CORI-) CORIXA CORP.  
 XX Xu J, Dillion DC;  
 XX WPI: 2001-440862/47.  
 XX P-PSDB; RAU04961.  
 XX Novel polynucleotide encoding polypeptide comprising a portion of  
 XX prostate tumour protein useful for inhibiting development of prostate  
 XX cancer or for treating prostate cancer in a patient  
 XX Claim 3; Column 118-121; 105pp; English.  
 XX The sequence is a human prostate tumour cDNA which encodes a  
 XX partial tumour protein. The DNA is useful for inhibiting the development  
 XX of prostate cancer or for treating prostate cancer in a patient.  
 XX Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;  
 SQ

Query Match 100.0%; Score 3409.6; DB 22; Length 3410;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGAAACACAGCTGACCGGCTGCGGCTGACAGCGCGCGCTCGGCCAGGATCTGA 60  
 Db 1 GGGAAACACAGCTGACCGGCTGCGGCTGACAGCGCGCGCTCGGCCAGGATCTGA 60  
 QY 61 GTGATGAGACGTGTCCCACTGAGTGCCCAACAGCAGCAGGTGTGAGCATGGGCTGAG 120  
 Db 61 GTGATGAGACGTGTCCCACTGAGTGCCCAACAGCAGCAGGTGTGAGCATGGGCTGAG 120  
 QY 121 AAGCTGGACCGCACAAAGGCTGGCAGAAATGGGCCCTGGCTGATTCCTAGGCAGTT 180  
 Db 121 AAGCTGGACCGCACAAAGGCTGGCAGAAATGGGCCCTGGCTGATTCCTAGGCAGTT 180  
 QY 181 GCGCGCACAGGAGGAGGAGGCGGCTTCTGGAGCAGCGCCGAGCAGGAGGAGGAGGAGGAGG 240

Db 181 GCGGCGACGAGGAGGAGCGCGACCTTCTCGAGCAGAGCGCGAGACGAGTTCG 240  
QY 241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCTGCCCCACTATGTCACAGAGGCTG 300  
Db 241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCTGCCCCACTATGTCACAGAGGCTG 300  
QY 301 GGTGAGCGGCTGCTGCGGACCGAAGCCAGCTTTCGTGGTCAACCTGCTAACCTT 360  
Db 301 GGTGAGCGGCTGCTGCGGACCGAAGCCAGCTTTCGTGGTCAACCTGCTAACCTT 360  
QY 361 TGGCCTGGAGGTGTGTTGGCCGAGGCATCACTATGTGCCGCCCTCTGCTGCGAAGT 420  
Db 361 TGGCCTGGAGGTGTGTTGGCCGAGGCATCACTATGTGCCGCCCTCTGCTGCGAAGT 420  
QY 421 GGGGGTAGAGGAAGTTCATGACCATTGCTGCGGCATTTGTCAGTCTGGGCTGCT 480  
Db 421 GGGGGTAGAGGAAGTTCATGACCATTGCTGCGGCATTTGTCAGTCTGGGCTGCT 480  
QY 481 CTGTGTCGCCCTCCTAGGCTCAGCCAGTGACCACTGGCGTGACGCTATGGCCGCGCG 540  
Db 481 CTGTGTCGCCCTCCTAGGCTCAGCCAGTGACCACTGGCGTGACGCTATGGCCGCGCG 540  
QY 541 GCCCTTCATCTGGGCACTGCTGTTGGCAATCCTGTGAGCCCTTTCTCATCCCAAGGC 600  
Db 541 GCCCTTCATCTGGGCACTGCTGTTGGCAATCCTGTGAGCCCTTTCTCATCCCAAGGC 600  
QY 601 CGGCTGGCTAGCAGGCTGCTGTCGCCGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
Db 601 CGGCTGGCTAGCAGGCTGCTGTCGCCGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
QY 661 CTTGGGCGTGGGCTGCTGGACTTCTGTGGCCAGGTGTCTTCACTCCACTGGAGGCCCT 720  
Db 661 CTTGGGCGTGGGCTGCTGGACTTCTGTGGCCAGGTGTCTTCACTCCACTGGAGGCCCT 720  
QY 721 GCTCTCTGACTCTTCCGGGACCCGGACCACTGTGCGCAGGCTACTCTGTCTATGCTT 780  
Db 721 GCTCTCTGACTCTTCCGGGACCCGGACCACTGTGCGCAGGCTACTCTGTCTATGCTT 780  
QY 781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 TGGCCTGGGCCCCCTACCTGGGACCCAGAGGAGTGCCTTCTTGGCCTGCTCACCCCTCAT 900  
Db 841 TGGCCTGGGCCCCCTACCTGGGACCCAGAGGAGTGCCTTCTTGGCCTGCTCACCCCTCAT 900  
QY 901 CTTCTCTACCTCGTAGCAGCACACTGCTGGTGGCTGAGGAGGCGCTGGGGCCAC 960  
Db 901 CTTCTCTACCTCGTAGCAGCACACTGCTGGTGGCTGAGGAGGCGCTGGGGCCAC 960  
QY 961 CGAGCAGCAGAGGCGCTGTCGGCCCCCTCTGTCGCCCCACTGCTGCTCCATGCCGGC 1020  
Db 961 CGAGCAGCAGAGGCGCTGTCGGCCCCCTCTGTCGCCCCACTGCTGCTCCATGCCGGC 1020  
QY 1021 CGGCTTGGCTTCCGGAACCTGGGGCCCTGCTTCCCGGCTGACCACTGCTGCTGCG 1080  
Db 1021 CGGCTTGGCTTCCGGAACCTGGGGCCCTGCTTCCCGGCTGACCACTGCTGCTGCG 1080  
QY 1081 CATGCCCGCACCTCGCCGCTCTGCTGGCTGAGCTGTCAGCTGGATGGGCACTCAT 1140  
Db 1081 CATGCCCGCACCTCGCCGCTCTGCTGGCTGAGCTGTCAGCTGGATGGGCACTCAT 1140  
QY 1141 GACCTTCACGCTGTTTTACGGATTTCTGTTGGGAGGGGCTGTACAGGCGCTGCCAG 1200  
Db 1141 GACCTTCACGCTGTTTTACGGATTTCTGTTGGGAGGGGCTGTACAGGCGCTGCCAG 1200  
QY 1201 AGCTGAGCGGGCAGCGAGGCGCGGACACTATGATGAAGGCTTCCGATGGGCACT 1260  
Db 1201 AGCTGAGCGGGCAGCGAGGCGCGGAGACACTATGATGAAGGCTTCCGATGGGCACT 1260  
QY 1261 GGGGCTGTTCTGAGTGGCCATCTCCCTGGCTCTCTCTGCTGCTATGACCGGCTGCT 1320  
Db 1261 GGGGCTGTTCTGAGTGGCCATCTCCCTGGCTCTCTCTGCTGCTATGACCGGCTGCT 1320

QY 1321 GCAGCGATTGGCACTCGAGCAGTCTATTGSCCACTGTGCAGCTTTCCTGTGGCTGC 1380  
Db 1321 GCAGCGATTGGCACTCGAGCAGTCTATTGSCCACTGTGCAGCTTTCCTGTGGCTGC 1380  
QY 1381 CGGTGCCACATGCTTGTCCACAGTGTGGCGTGTGACAGCTTTCAGCGGCTCACCGG 1440  
Db 1381 CGGTGCCACATGCTTGTCCACAGTGTGGCGTGTGACAGCTTTCAGCGGCTCACCGG 1440  
QY 1441 GTTCACCTTCTCAGCCCTGCGAGATTCCTCCCTACACACTGAGCTTACACCGGGA 1500  
Db 1441 GTTCACCTTCTCAGCCCTGCGAGATTCCTCCCTACACACTGAGCTTACACCGGGA 1500  
QY 1501 GAAGCAGGTGTTCCTGCCCCAATACCGAGGGGACACTGAGGTGCTAGCAGTGGAGAC 1560  
Db 1501 GAAGCAGGTGTTCCTGCCCCAATACCGAGGGGACACTGAGGTGCTAGCAGTGGAGAC 1560  
QY 1561 CCTGATGACAGCTTCTGCCCCAGGCTTAAGCTTGGAGTCCCTTCCCTAATGACACGT 1620  
Db 1561 CCTGATGACAGCTTCTGCCCCAGGCTTAAGCTTGGAGTCCCTTCCCTAATGACACGT 1620  
QY 1621 GGGTGTGGAGGAGTGGCTGCTCCACCTCCACCGCTGCGGGCTGCGGGCTGCGCTG 1680  
Db 1621 GGGTGTGGAGGAGTGGCTGCTCCACCTCCACCGCTGCGGGCTGCGGGCTGCGCTG 1680  
QY 1681 TGATGTCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
Db 1681 TGATGTCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
QY 1741 GGGGATCTGCTGACCTCGCCATCTCCACCTCCACCGCTGCGGGCTGCGGGCTGCGCTG 1800  
Db 1741 GGGGATCTGCTGACCTCGCCATCTCCACCTCCACCGCTGCGGGCTGCGGGCTGCGCTG 1800  
QY 1801 ATCCCTGTTATGGCTCCATTGTCAGCTCAGCAGTCTGCTCACTGCTATATGCTGTC 1860  
Db 1801 ATCCCTGTTATGGCTCCATTGTCAGCTCAGCAGTCTGCTCACTGCTATATGCTGTC 1860  
QY 1861 TGGCGAGGCTGGGTCTGGTGGCACTTACTTGTGTACACAGTGTATTTGACAAAG 1920  
Db 1861 TGGCGAGGCTGGGTCTGGTGGCACTTACTTGTGTACACAGTGTATTTGACAAAG 1920  
QY 1921 CGACTTGGCCAAATACCTCAGCTAGAACTTCCACCACTTGGGTGGAGGCTGCTGCT 1980  
Db 1921 CGACTTGGCCAAATACCTCAGCTAGAACTTCCACCACTTGGGTGGAGGCTGCTGCT 1980  
QY 1981 CACTGGTCCAGCTCCCCCTCTCTGTAGCCCATGGGCTGCGGGCTGGCGGCTGCT 2040  
Db 1981 CACTGGTCCAGCTCCCCCTCTCTGTAGCCCATGGGCTGCGGGCTGGCGGCTGCT 2040  
QY 2041 TTTCTGTGCTGCCAAAGTAAATGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100  
Db 2041 TTTCTGTGCTGCCAAAGTAAATGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100  
QY 2101 GCTGACAGCTGGGGCTGGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160  
Db 2101 GCTGACAGCTGGGGCTGGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160  
QY 2161 ACTGGAGGCTTCCAGAGGCTTTCAGTGTGAGTATATACAGGAGGCGCCAGAGGCTCC 2220  
Db 2161 ACTGGAGGCTTCCAGAGGCTTTCAGTGTGAGTATATACAGGAGGCGCCAGAGGCTCC 2220  
QY 2221 ATGCACTGGAATCGGGAGCTGTCAGGTGGATTTACCCAGGCTCAGGGTTACAGCTAGC 2280  
Db 2221 ATGCACTGGAATCGGGAGCTGTCAGGTGGATTTACCCAGGCTCAGGGTTACAGCTAGC 2280  
QY 2281 CTCTAGTTGAGACACACCTTAGAAGGCTTTTGGAGCTGAATAAAGCTCAGTACCTG 2340  
Db 2281 CTCTAGTTGAGACACACCTTAGAAGGCTTTTGGAGCTGAATAAAGCTCAGTACCTG 2340  
QY 2341 GTTTCCTCTCTAACGCTTAACTGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400  
Db 2341 GTTTCCTCTCTAACGCTTAACTGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400



QY 2401 TTCTAGGATGAACACACTCTCCATGGGATTTGAACATATGACTTATTTAGGGGAAGA 2460  
 DB |||||||  
 QY 2401 TTCTAGGATGAACACACTCTCCATGGGATTTGAACATATGACTTATTTAGGGGAAGA 2460  
 DB |||||||  
 QY 2461 GTCTGAGGGGCAACACAGAACAGAGTCCCTCAGCCACACAGCTGTCTTTTGGCT 2520  
 DB |||||||  
 QY 2461 GTCTGAGGGGCAACACAGAACAGAGTCCCTCAGCCACACAGCTGTCTTTTGGCT 2520  
 DB |||||||  
 QY 2521 GATCCACCCCTCTTACCTTTATCAGGATGTGGCTGTGGTCTCTCTGTGCCATCA 2580  
 DB |||||||  
 QY 2521 GATCCACCCCTCTTACCTTTATCAGGATGTGGCTGTGGTCTCTCTGTGCCATCA 2580  
 DB |||||||  
 QY 2581 CAGACACAGCAGCATTTAAATTTAACTTATTTAAACAAAGTAGAAGGGAATCCAT 2640  
 DB |||||||  
 QY 2581 CAGACACAGCAGCATTTAAATTTAACTTATTTAAACAAAGTAGAAGGGAATCCAT 2640  
 DB |||||||  
 QY 2641 TGCTAGCTTTCTGTGGTGTCTTAATTTATTTGGGTAGGTGGGGATCCCCACAATCA 2700  
 DB |||||||  
 QY 2641 TGCTAGCTTTCTGTGGTGTCTTAATTTATTTGGGTAGGTGGGGATCCCCACAATCA 2700  
 DB |||||||  
 QY 2701 GGTCCCTCAGATAGTGTCTATTTGGCTGATCTTGCAGAACTCTCTCTCCGCGGT 2760  
 DB |||||||  
 QY 2701 GGTCCCTCAGATAGTGTCTATTTGGCTGATCTTGCAGAACTCTCTCTCCGCGGT 2760  
 DB |||||||  
 QY 2761 CTGGCCCCCAAAATGCCCTAACCCAGGACCTTTGAAATTTCTACTCATCCCAATGATAAT 2820  
 DB |||||||  
 QY 2761 CTGGCCCCCAAAATGCCCTAACCCAGGACCTTTGAAATTTCTACTCATCCCAATGATAAT 2820  
 DB |||||||  
 QY 2821 TCCAAATGCTGTACCCAGGTTAGGTGTGAAGGAAGGTAGAGGTGGGGCTTCAGGT 2880  
 DB |||||||  
 QY 2821 TCCAAATGCTGTGTACCCAGGTTAGGTGTGAAGGAAGGTAGAGGTGGGGCTTCAGGT 2880  
 DB |||||||  
 QY 2881 CTCACGGCTTCCCTAACCCACCCCTCTCTCTTTGGCCAGCCTGTCTCCCCCACTTCCA 2940  
 DB |||||||  
 QY 2881 CTCACGGCTTCCCTAACCCACCCCTCTCTCTTTGGCCAGCCTGTCTCCCCCACTTCCA 2940  
 DB |||||||  
 QY 2941 CTCCTCTTACTCTCTTAGGACTGGCTGATGAAGGCACCTGCCCAAAATTTCCCTACC 3000  
 DB |||||||  
 QY 2941 CTCCTCTTACTCTCTTAGGACTGGCTGATGAAGGCACCTGCCCAAAATTTCCCTACC 3000  
 DB |||||||  
 QY 3001 CCCACTTTCCTTACCCCACTTCCCAACAGCTCCCAACCCCTGTCTTGGAGCTACT 3060  
 DB |||||||  
 QY 3001 CCCACTTTCCTTACCCCACTTCCCAACAGCTCCCAACCCCTGTCTTGGAGCTACT 3060  
 DB |||||||  
 QY 3061 GCAGGACCAAGAGCAAAAGTCGGCTTCCCAAGCCTTTGTCTCATCTCAGCCCCAGAGT 3120  
 DB |||||||  
 QY 3061 GCAGGACCAAGAGCAAAAGTCGGCTTCCCAAGCCTTTGTCTCATCTCAGCCCCAGAGT 3120  
 DB |||||||  
 QY 3121 ATATCTGTCTTGGGAATCTCACACAGAACTCAGGACACCCCTGCCTGAGCTAAGG 3180  
 DB |||||||  
 QY 3121 ATATCTGTCTTGGGAATCTCACACAGAACTCAGGACACCCCTGCCTGAGCTAAGG 3180  
 DB |||||||  
 QY 3181 GAGGTCTTATCTCAGGGGGGTTTAAAGTCCCGTTTGAATATGTCTTATTATT 3240  
 DB |||||||  
 QY 3181 GAGGTCTTATCTCAGGGGGGTTTAAAGTCCCGTTTGAATATGTCTTATTATT 3240  
 DB |||||||  
 QY 3241 TAGCGGGGTGAATTTTATCTGTAAAGTGAGCAATCAGAGTATATGTTATGTGTACA 3300  
 DB |||||||  
 QY 3241 TAGCGGGGTGAATTTTATCTGTAAAGTGAGCAATCAGAGTATATGTTATGTGTACA 3300  
 DB |||||||  
 QY 3301 AAATTAAGGCTTCTTATATGTTTAAAAAAAATAAAAAAATAAAAAAATAAAAAA 3360  
 DB |||||||  
 QY 3301 AAATTAAGGCTTCTTATATGTTTAAAAAAAATAAAAAAATAAAAAAATAAAAAA 3360  
 DB |||||||  
 QY 3361 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 3410  
 DB |||||||  
 QY 3361 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 3410  
 DB |||||||

RESULT 6

AAH93465

ID AAH93465 standard; cDNA; 3410 BP.

XX

AAH93465;

04-OCT-2001 (first entry)

Human prostate-specific full length cDNA sequence LI-12.

Human; prostate cancer; prostate-specific; diagnosis; vaccine; cytostatic; gene therapy; metastasis; ss.

Homo sapiens.

WO200151633-A2.

19-JUL-2001.

16-JAN-2001; 2001WO-US01574.

14-JAN-2000; 2000US-0483672.

(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;  
 Wang A, Meagher MJ;  
 WPI; 2001-425873/45.

New polynucleotide encoding a prostate-specific protein, for  
 diagnosing, monitoring and treating prostate cancer in a patient and  
 for use in vaccines -

Claim 1; Page 265-266; 543pp; English.

The present invention describes polynucleotide sequences (I) which encode  
 prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
 and can be used in vaccine production and gene therapy. (I), (II),  
 antibodies to (II), fusion proteins comprising (II), and isolated  
 T cells prepared using (I) or (II) are used to treat cancer in a patient.  
 (I) and the antibodies are also used in the detection of cancer in a  
 patient. The cancer that is diagnosed or treated is particularly  
 prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
 (I) can be used for monitoring the progression of cancer in a patient.  
 (I) and (II) can also be used to improve diagnostic and therapeutic  
 methods for prostate cancer. They can indicate the level of metastasis  
 as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to  
 AAH01318 represent polynucleotide and amino acid sequences used in the  
 exemplification of the present invention.

Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

Query Match 100.0%; Score 3409.6; DB 22; Length 3410;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGAAACAGCCTGCACCGCTGGCTCCGGGTGACAGCCGCGCTCGGCAGGATCTGA 60  
 |||||||

DB 1 GGGAAACAGCCTGCACCGCTGGCTCCGGGTGACAGCCGCGCTCGGCAGGATCTGA 60  
 |||||||

QY 61 GTGATGAGAGCTGTCCCACTAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120  
 |||||||

DB 61 GTGATGAGAGCTGTCCCACTAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120  
 |||||||

QY 121 AAGCTGGACCGGACCAAGGCTGGCAGAAATGGCGCTGATTCCTAGGCAGTT 180  
 |||||||

DB 121 AAGCTGGACCGGACCAAGGCTGGCAGAAATGGCGCTGATTCCTAGGCAGTT 180  
 |||||||

QY 181 GCGCGCAGCAGGAGAGGCGCCAGCTTCTGTGAGCAGAGCCGAGCAGAGCTTCTG 240  
 |||||||

DB 181 GCGCGCAGCAGGAGAGGCGCCAGCTTCTGTGAGCAGAGCCGAGCAGAGCTTCTG 240  
 |||||||

QY 241 GAGTGCCTGAACGCGCCCTGAGCCCTACCGCTGAGCCCTGAGCCCTGAGCCCTGAG 300  
 |||||||

DB 241 GAGTGCCTGAACGCGCCCTGAGCCCTACCGCTGAGCCCTGAGCCCTGAGCCCTGAG 300  
 |||||||



QY 301 GGTGAGCGCCTGCTGGGACCGGAAGCCAGCTCTTGTGCTCAACCTGCTAACCTT 360  
DB 301 GGTGAGCGCCTGCTGGGACCGGAAGCCAGCTCTTGTGCTCAACCTGCTAACCTT 360  
QY 361 TGGCCTGAGAGTGTTTGGCCGAGGCATCACTATGTGCCGCCCTCTGCTGTGGAAGT 420  
DB 361 TGGCCTGAGAGTGTTTGGCCGAGGCATCACTATGTGCCGCCCTCTGCTGTGGAAGT 420  
QY 421 GGGGTAGAGAGAAAGTTATGACCATGGTGTGGGCAATGGTCCAGTGTGGGCCCTGGT 480  
DB 421 GGGGTAGAGAGAAAGTTATGACCATGGTGTGGGCAATGGTCCAGTGTGGGCCCTGGT 480  
QY 481 CTGTGTCGGCTCTAGGCTCAGCCAGTGAACCTGCGCTGGAGCTATGGCCGCCGCCG 540  
DB 481 CTGTGTCGGCTCTAGGCTCAGCCAGTGAACCTGCGCTGGAGCTATGGCCGCCGCCG 540  
QY 541 GCCCTTCACTGAGGCACTGCTTGGGCACTCTGCTGAGCCTCTTCTCATCCCAAGGC 600  
DB 541 GCCCTTCACTGAGGCACTGCTTGGGCACTCTGCTGAGCCTCTTCTCATCCCAAGGC 600  
QY 601 CGGCTGCTAGCAGGCTGCTGTGCCGGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
DB 601 CGGCTGCTAGCAGGCTGCTGTGCCGGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
QY 661 CCTGGGCGTGGGCTGCTGAGCTTCTGTGGCAGGTGCTTCACTCCACTGGAGGCCCT 720  
DB 661 CCTGGGCGTGGGCTGCTGAGCTTCTGTGGCAGGTGCTTCACTCCACTGGAGGCCCT 720  
QY 721 GCTCTGACCTCTTCGGGACCGGACCACTGTGCGCAGGCACTGCTGCTATGCCCT 780  
DB 721 GCTCTGACCTCTTCGGGACCGGACCACTGTGCGCAGGCACTGCTGCTATGCCCT 780  
QY 781 CATGATCAGTCTTGGGCTGCTGAGCTACCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 781 CATGATCAGTCTTGGGCTGCTGAGCTACCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 TGCCCTGGCCCTTACCTGGGACCGGAGGAGTGCCTTTTGGCCTGCTCAACCTCAT 900  
DB 841 TGCCCTGGCCCTTACCTGGGACCGGAGGAGTGCCTTTTGGCCTGCTCAACCTCAT 900  
QY 901 CTTCTCACTGCTAGCAGCACACTGCTGGTGGCTGAGGAGCAGCGCTGGGCCCCAC 960  
DB 901 CTTCTCACTGCTAGCAGCACACTGCTGGTGGCTGAGGAGCAGCGCTGGGCCCCAC 960  
QY 961 CGAGCCAGCAGAGGCTGTGGGCCCTCTTGTGGCCCACTGCTGCTGCTGCTGCTGCT 1020  
DB 961 CGAGCCAGCAGAGGCTGTGGGCCCTCTTGTGGCCCACTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 CCGCTTGGCTTCCGGAACTGGGCGCCCTGCTTCCCGGCTGCACCACTGCTGCTGCT 1080  
DB 1021 CCGCTTGGCTTCCGGAACTGGGCGCCCTGCTTCCCGGCTGCACCACTGCTGCTGCT 1080  
QY 1081 CATGCCCGCACTGCTGGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
DB 1081 CATGCCCGCACTGCTGGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
QY 1141 GACCTTCACTGCTTGTACAGGATTTGCTGGGCGAGGCTGTACAGGCGCTGCCCCAG 1200  
DB 1141 GACCTTCACTGCTTGTACAGGATTTGCTGGGCGAGGCTGTACAGGCGCTGCCCCAG 1200  
QY 1201 AGCTGAGCCGGGACCGGAGCCGAGACACTATGATGAAGGCTTCCGATGGGCACTT 1260  
DB 1201 AGCTGAGCCGGGACCGGAGCCGAGACACTATGATGAAGGCTTCCGATGGGCACTT 1260  
QY 1261 GGGCTGTTCTGAGTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
DB 1261 GGGCTGTTCTGAGTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 CGAGGATTCGGCACTCGAGAGTCTATTTGGCCAGTGTGGCACTTTCCTGTGGCTGC 1380  
DB 1321 CGAGGATTCGGCACTCGAGAGTCTATTTGGCCAGTGTGGCACTTTCCTGTGGCTGC 1380

QY 1381 CGGTGCCACATGCTTCCACAGTGTGGCGTGTGACAGCTTCCAGCGGCCCTCACCGG 1440  
DB 1381 CGGTGCCACATGCTTCCACAGTGTGGCGTGTGACAGCTTCCAGCGGCCCTCACCGG 1440  
QY 1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTTACACACTGGCCCTCCCTCTACCCACGGGA 1500  
DB 1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTTACACACTGGCCCTCCCTCTACCCACGGGA 1500  
QY 1501 GAAGCAGTGTTCTGCCCAAATACCGAGGGGACACTGGAGGTGTAGCAGTGAAGACAG 1560  
DB 1501 GAAGCAGTGTTCTGCCCAAATACCGAGGGGACACTGGAGGTGTAGCAGTGAAGACAG 1560  
QY 1561 CCTGATGACCACTTCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTTAATGGACAGT 1620  
DB 1561 CCTGATGACCACTTCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTTAATGGACAGT 1620  
QY 1621 GGGTGTGAGGACAGTGGCCTGCTCCCACTCCACCCCGCTCTCGCGGGCTCTGCGCTG 1680  
DB 1621 GGGTGTGAGGACAGTGGCCTGCTCCCACTCCACCCCGCTCTCGCGGGCTCTGCGCTG 1680  
QY 1681 TGATGCTCCGTACGTGTGGTGGTGGTGAGCCCAACGAGGCCAGGGTGTTCGGGGCG 1740  
DB 1681 TGATGCTCCGTACGTGTGGTGGTGGTGAGCCCAACGAGGCCAGGGTGTTCGGGGCG 1740  
QY 1741 GGGCATCTGCCCTGGACCTGCCATCCTGGATAGTGCCTTCTGCTGCCAGTGGCCCC 1800  
DB 1741 GGGCATCTGCCCTGGACCTGCCATCCTGGATAGTGCCTTCTGCTGCCAGTGGCCCC 1800  
QY 1801 ATCCCTGTTATAGGGCTCCATTTGCTACAGCTACGAGCTGTCTACCTGCTATATGGTGT 1860  
DB 1801 ATCCCTGTTATAGGGCTCCATTTGCTACAGCTACGAGCTGTCTACCTGCTATATGGTGT 1860  
QY 1861 TGCCGAGCCTTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1920  
DB 1861 TGCCGAGCCTTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1920  
QY 1921 CGACTTGGCCAAATACTCAGCGTGAACACTTCCAGCACATTTGGGGTGGAGGCCCTGCT 1980  
DB 1921 CGACTTGGCCAAATACTCAGCGTGAACACTTCCAGCACATTTGGGGTGGAGGCCCTGCT 1980  
QY 1981 CACTGGGTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
DB 1981 CACTGGGTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
QY 2041 TTCTGTGCTGCAAGTAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2100  
DB 2041 TTCTGTGCTGCAAGTAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2100  
QY 2101 GCTGCACAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 2160  
DB 2101 GCTGCACAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 2160  
QY 2161 ACTGGAGGCTTCCAAAGGGGTTTTCAGTGTGAGCTTATACAGGAGGCCAGAGGGCTCC 2220  
DB 2161 ACTGGAGGCTTCCAAAGGGGTTTTCAGTGTGAGCTTATACAGGAGGCCAGAGGGCTCC 2220  
QY 2221 ATGCACTGGAATTCGGGGACTCTGCAAGTGGATTTACCAGGCTCAGGGTTAACAGCTAGC 2280  
DB 2221 ATGCACTGGAATTCGGGGACTCTGCAAGTGGATTTACCAGGCTCAGGGTTAACAGCTAGC 2280  
QY 2281 CTCTAGTTGAGACACACCTAGAGAGGTTTTCGGAGCTGAATAAATCACTAGTCACCTG 2340  
DB 2281 CTCTAGTTGAGACACACCTAGAGAGGTTTTCGGAGCTGAATAAATCACTAGTCACCTG 2340  
QY 2341 GTTTCCTCTCTAAGCCCTTTAAGCTGAGCTTCTGCTTAAATGAGCTTTCATGGGAG 2400  
DB 2341 GTTTCCTCTCTAAGCCCTTTAAGCTGAGCTTCTGCTTAAATGAGCTTTCATGGGAG 2400  
QY 2401 TTCTAGGATGAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460  
DB 2401 TTCTAGGATGAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460  
QY 2461 GTCTTGGGGCAACACACAGAACAGGTCCTCCCTCAGGCCACAGCACTGTCTTTTGTCT 2520





Query Match	100.0%;	Score 3409.6;	DB 22;	Length 3410;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3410;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGGAACAGCCTGCACGCGCTGGCTCGGGGTGACAGCGCGCGCTCGGCACGAGATCTGA	60	
DB	1	GGGAACAGCCTGCACGCGCTGGCTCGGGGTGACAGCGCGCGCTCGGCACGAGATCTGA	60	
QY	61	GTGATGAGAGCTGTCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG	120	
DB	61	GTGATGAGAGCTGTCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG	120	
QY	121	AAGCTGAGCCGGCAACAAAGGGCTGGCAGAAATGGGGCCCTGGCTGATTCCTAGGCAAGTT	180	
DB	121	AAGCTGAGCCGGCAACAAAGGGCTGGCAGAAATGGGGCCCTGGCTGATTCCTAGGCAAGTT	180	
QY	181	GGCGGCAGCAGGAGGAGGCGCGCACTTCTGGAGCAGACGCCAGACGAAGCAGATTCTG	240	
DB	181	GGCGGCAGCAGGAGGAGGCGCGCACTTCTGGAGCAGACGCCAGACGAAGCAGATTCTG	240	
QY	241	GAGTGCCTGAACGGCCCGCTGAGCCCTACCGCCTGGCCCACTATGCTCCAGAGGCTGTG	300	
DB	241	GAGTGCCTGAACGGCCCGCTGAGCCCTACCGCCTGGCCCACTATGCTCCAGAGGCTGTG	300	
QY	301	GGTGAAGCCGCTGCTGGCGCACCGGAAGCCCAAGCTCTTGGTGTGCTCAACCTGTAACTTT	360	
DB	301	GGTGAAGCCGCTGCTGGCGCACCGGAAGCCCAAGCTCTTGGTGTGCTCAACCTGTAACTTT	360	
QY	361	TGGCCTGGAGGTGTGTTTGGCCGCGAGGCATCACCTATGTGCCGCTCTGCTGCTGGAGTT	420	
DB	361	TGGCCTGGAGGTGTGTTTGGCCGCGAGGCATCACCTATGTGCCGCTCTGCTGCTGGAGTT	420	
QY	421	GGGGGTAGAGGAGAAGTTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCCTGGGCTGGT	480	
DB	421	GGGGGTAGAGGAGAAGTTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCCTGGGCTGGT	480	
QY	481	CTGTGTCCCGCTCCTAGGCTCAGGCAGTGAACCTAGGCTGGCGTGGACGCTATGGCCGCCGCCG	540	

RESULT 8	
AAH02530	
ID	AAH02530 standard; cDNA; 3410 BP.
XX	
XX	
AC	AAH02530;
XX	
XX	
DT	14-JUN-2001 (first entry)
XX	
DE	Prostate tumour antigen determined full length cDNA sequence for LI-12.
XX	
KW	Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW	prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX	
OS	Homo sapiens.
XX	

|||||  
Db 481 CTGTGTCCTCCCTAGGCTAGCCAGTACCACATGGCGTGACGCTATGGCCGCGCG 540  
QY 541 GCCCTTCATCTGGGCACTGCTCTGGGCACTCCTGTGTGAGCTCTTCTTCATCCCAAGGC 600  
Db 541 GCCCTTCATCTGGGCACTGCTCTGGGCACTCCTGTGTGAGCTCTTCTTCATCCCAAGGC 600  
QY 601 CGGCTGGCTAGCAGGCTGCTGTGCCCGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
Db 601 CGGCTGGCTAGCAGGCTGCTGTGCCCGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
QY 661 CTGGGCTGGGCTGCTGGACTTCTGTGCCAGGTGTCTCACTCCACATGGAGGCCCT 720  
Db 661 CTGGGCTGGGCTGCTGGACTTCTGTGCCAGGTGTCTCACTCCACATGGAGGCCCT 720  
QY 721 GCTCTGACCTCTTCGGGACCCGACACATGTGCCAGGCTACTCTGTCTATGGCTT 780  
Db 721 GCTCTGACCTCTTCGGGACCCGACACATGTGCCAGGCTACTCTGTCTATGGCTT 780  
QY 781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 TGGCTTGGCCCTTACCTGGGACCCAGAGGAGTCTCTTTGGGCTGCTCACTGCTCAT 900  
Db 841 TGGCTTGGCCCTTACCTGGGACCCAGAGGAGTCTCTTTGGGCTGCTCACTGCTCAT 900  
QY 901 CTTCTCAGCTGCTAGCAGCCACACTGCTGTGTGCTGAGAGGAGCGCTGGGCCAC 960  
Db 901 CTTCTCAGCTGCTAGCAGCCACACTGCTGTGTGCTGAGAGGAGCGCTGGGCCAC 960  
QY 961 CGAGCCAGCAGAGGCTGTGGGCCCCCTCTGTGTCGCCCACTGCTGCTGCTGCTGCTGCT 1020  
Db 961 CGAGCCAGCAGAGGCTGTGGGCCCCCTCTGTGTCGCCCACTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 CGGCTTGGCTTTCGGGAACCTGGGCGCCCTGCTTCCCGGCTGACACAGCTGTGCTGCG 1080  
Db 1021 CGGCTTGGCTTTCGGGAACCTGGGCGCCCTGCTTCCCGGCTGACACAGCTGTGCTGCG 1080  
QY 1081 CATGCCCGCACCTTGGCCGCTCTGTGGCTGAGCTGTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 1081 CATGCCCGCACCTTGGCCGCTCTGTGGCTGAGCTGTGCTGCTGCTGCTGCTGCTGCT 1140  
QY 1141 GACCTTCAGCTGTGTTACAGGATTTCTGGGCGAGGCTGTACAGGCGCTGCCAG 1200  
Db 1141 GACCTTCAGCTGTGTTACAGGATTTCTGGGCGAGGCTGTACAGGCGCTGCCAG 1200  
QY 1201 AGCTGAGCGGCGACCGGCGGAGACACTATGATGAAGCGCTTCGGATGGGAGCCT 1260  
Db 1201 AGCTGAGCGGCGACCGGCGGAGACACTATGATGAAGCGCTTCGGATGGGAGCCT 1260  
QY 1261 GGGGCTGTCTGACAGTGGCCATCTCCCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 1320  
Db 1261 GGGGCTGTCTGACAGTGGCCATCTCCCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 GCAGGATTCGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCTGTGGCTGC 1380  
Db 1321 GCAGGATTCGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCTGTGGCTGC 1380  
QY 1381 CGGTGCCACATGCTGTCCTCCACAGTGTGGCGTGGTACAGCTTCAGCGGCCCTCACCG 1440  
Db 1381 CGGTGCCACATGCTGTCCTCCACAGTGTGGCGTGGTACAGCTTCAGCGGCCCTCACCG 1440  
QY 1441 GFTTACCTTCAGCCCTTCAGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
Db 1441 GFTTACCTTCAGCCCTTCAGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
QY 1501 GAAGCAGGTGTCTGTCCTCCAAATACCGAGGGACACTGGAGGTGCTAGCAGTAGGACAG 1560  
Db 1501 GAAGCAGGTGTCTGTCCTCCAAATACCGAGGGACACTGGAGGTGCTAGCAGTAGGACAG 1560  
QY 1561 CCGTATGACAGCTTCTGTCGAGCGCCCTAGCCCTGGAGCTCCCTTCCCTAATGACACGT 1620  
|||||

Db 1561 CCTGATGACCAGCTTCTGTCGAGGCCCTAAGCCCTGGAGCTCCCTTCCCTAATGACACGT 1620  
QY 1621 GGGTCTGGAGCAGTGGCTGCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCGCTG 1680  
Db 1621 GGGTCTGGAGCAGTGGCTGCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCGCTG 1680  
QY 1681 TGATGCTCCGTAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
Db 1681 TGATGCTCCGTAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
QY 1741 GGGCATCTGCTGACCTCGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
Db 1741 GGGCATCTGCTGACCTCGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
QY 1801 ATCCCTGTTATGGCTCCCATTTGTCAGCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
Db 1801 ATCCCTGTTATGGCTCCCATTTGTCAGCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
QY 1861 TGCCGAGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
Db 1861 TGCCGAGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
QY 1921 CGACTTGGCCAAATACCTCAGCTAGAAAACCTCCAGACATTTGGGTTGGAGGCTGCTGCT 1980  
Db 1921 CGACTTGGCCAAATACCTCAGCTAGAAAACCTCCAGACATTTGGGTTGGAGGCTGCTGCT 1980  
QY 1981 CACTGGTCCAGCTCCCGCTCCTGTTAGCCCATGGGCTGCGGGCTGGCGGCTGCTGCTGCT 2040  
Db 1981 CACTGGTCCAGCTCCCGCTCCTGTTAGCCCATGGGCTGCGGGCTGGCGGCTGCTGCTGCT 2040  
QY 2041 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
Db 2041 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
QY 2101 GCTGACAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
Db 2101 GCTGACAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
QY 2161 ACTGGAGGCTTCCAGAGGCTTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
Db 2161 ACTGGAGGCTTCCAGAGGCTTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
QY 2221 ATGCACTGGAATCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Db 2221 ATGCACTGGAATCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
QY 2281 CTCTAGTTGAGACACACCTTAGAAGGCTTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
Db 2281 CTCTAGTTGAGACACACCTTAGAAGGCTTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
QY 2341 GTTTCCTGCTTAAAGCCCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
Db 2341 GTTTCCTGCTTAAAGCCCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
QY 2401 TTTCTAGGATGAACACTCTCCATGGATTTGAACATATGACTTATTTGTTAGGGAAGA 2460  
Db 2401 TTTCTAGGATGAACACTCTCCATGGATTTGAACATATGACTTATTTGTTAGGGAAGA 2460  
QY 2461 GTCTGAGGGGCAACACAGGCTCCCTCAGCCCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
Db 2461 GTCTGAGGGGCAACACAGGCTCCCTCAGCCCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
QY 2521 GATCCACCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580  
Db 2521 GATCCACCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580  
QY 2581 CAGACACAGGCTTAAATATTTAACTTATTTAACTTATTTAACTTATTTAACTTATTTAACTTATTT 2640  
Db 2581 CAGACACAGGCTTAAATATTTAACTTATTTAACTTATTTAACTTATTTAACTTATTTAACTTATTT 2640  
QY 2641 TGTAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
Db 2641 TGTAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700

```
QY 2701 GGTCCCTGAGATAGCTGGTCAATGGGCTGATCATTCGCCAGAAATCTTCTCTCTGGGT 2760
DB 2701 GGTCCCTGAGATAGCTGGTCAATGGGCTGATCATTCGCCAGAAATCTTCTCTCTGGGT 2760
QY 2761 CTGGCCCCCAAAATGCTTAACCCAGGACCTTGGAAATCTTACTCATCCCAATGATAAT 2820
DB 2761 CTGGCCCCCAAAATGCTTAACCCAGGACCTTGGAAATCTTACTCATCCCAATGATAAT 2820
QY 2821 TCCAAATGCTGTACCCCAAGGTTAGGCTGTTGAAGGAAGGTAGAGGTTGGGCTTCAGGT 2880
DB 2821 TCCAAATGCTGTACCCCAAGGTTAGGCTGTTGAAGGAAGGTAGAGGTTGGGCTTCAGGT 2880
QY 2881 CTCACAGGCTTCCCTTAACCAACCCCTCTTCTTGGCCAGCCTGGTTCGCCCACTTCCA 2940
DB 2881 CTCACAGGCTTCCCTTAACCAACCCCTCTTCTTGGCCAGCCTGGTTCGCCCACTTCCA 2940
QY 2941 CTCCCTCTACTCTCTAGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCTACG 3000
DB 2941 CTCCCTCTACTCTCTAGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCTACG 3000
QY 3001 CCCAACTTCCCTTACCCCAACTTTCCCAACAGCTCCCAACCCCTGTTGGAGTACT 3060
DB 3001 CCCAACTTCCCTTACCCCAACTTTCCCAACAGCTCCCAACCCCTGTTGGAGTACT 3060
QY 3061 GCAGGACCAAGCAACAAGTGGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
DB 3061 GCAGGACCAAGCAACAAGTGGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
QY 3121 ATATCTGTCTGGGATCTCACAGAACTCAGAGACCCCTGCTGAGCTAAGG 3180
DB 3121 ATATCTGTCTGGGATCTCACAGAACTCAGAGACCCCTGCTGAGCTAAGG 3180
QY 3181 GAGTCTTATCTCTCAGGGGGGTTAAGTGCCGTTTGAATAATCTCTCTATTATT 3240
DB 3181 GAGTCTTATCTCTCAGGGGGGTTAAGTGCCGTTTGAATAATCTCTCTATTATT 3240
QY 3241 TAGCGGGTGAATATTTTATCTGTAAGTGAGCAATCAGAGTATAATGTTTATGTTGACA 3300
DB 3241 TAGCGGGTGAATATTTTATCTGTAAGTGAGCAATCAGAGTATAATGTTTATGTTGACA 3300
QY 3301 AATTAAGGCTTCTTATATGTTTAAAAAATAAATAAATAAATAAATAAATAAATAA 3360
DB 3301 AATTAAGGCTTCTTATATGTTTAAAAAATAAATAAATAAATAAATAAATAAATAA 3360
QY 3361 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3410
DB 3361 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3410
```

## RESULT 9

ABL94929  
ID ABL94929 standard; cDNA; 3410 BP.

XX ABL94929;

XX 19-JUL-2002 (first entry)

XX Human L1-12 cDNA sequence SEQ ID NO 110.

XX Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
XX gene therapy; gene; ss.

XX Homo sapiens.

XX US200202248-A1.

XX 21-FEB-2002.

XX 12-JAN-2001; 2001US-0759143.

XX 25-FEB-1997; 97US-0806099.

PR 01-AUG-1997; 97US-0904804.

```
PR 09-FEB-1998; 98US-0020956.  
PR 25-FEB-1998; 98US-0030607.  
PR 14-JUL-1998; 98US-0115453.  
PR 23-SEP-1998; 98US-0159812.  
PR 15-JAN-1999; 99US-0232149.  
PR 09-APR-1999; 99US-0288946.  
PR 13-JUL-1999; 99US-0352616.  
PR 12-NOV-1999; 99US-0439313.  
PR 18-NOV-1999; 99US-0443686.  
PR 14-JAN-2000; 2000US-0483672.  
PR 27-MAR-2000; 2000US-0538857.  
PR 09-MAY-2000; 2000US-0568100.  
PR 12-MAY-2000; 2000US-0570737.  
PR 13-JUN-2000; 2000US-0593793.  
PR 27-JUN-2000; 2000US-0605783.  
PR 10-AUG-2000; 2000US-0636215.  
PR 29-AUG-2000; 2000US-0651236.  
PR 06-SEP-2000; 2000US-0652729.  
PR 02-OCT-2000; 2000US-0679426.  
PR 10-OCT-2000; 2000US-0685166.
```

XX (XUJ/) XU J.

PA (DILL/) DILLON D C.

PA (MITC/) MITCHAM J L.

PA (HARL/) HARLOCKER S L.

PA (JIAN/) JIANG Y.

PA (KALO/) KALOS M D.

PA (FANG/) FANGER G R.

PA (RETT/) RETTER M W.

PA (STOL/) STOLK J A.

PA (DAYC/) DAY C H.

PA (VEDV/) VEDVICK T S.

PA (CART/) CARTER D.

PA (LISX/) LI S X.

PA (WANG/) WANG A.

PA (SKEI/) SKEIKY Y A W.

PA (HEPL/) HEPLER W T.

PA (HEND/) HENDERSON R A.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;

PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;

PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2002-255649/30.

XX New prostate-specific polynucleotides for diagnosing and treating

PT diseases, in particular prostate cancer, and as markers for the

PT progression of cancer

XX Claim 1: SEQ ID NO 110; 87pp; English.

PS The present invention provides prostate-specific coding sequences and

XX their encoded proteins. These can be used in the diagnosis and treatment

CC of cancers, particularly prostate cancer. The present sequence is a cDNA

CC described in the invention.

XX Query Match

XX Best Local Similarity 100.0%; Score 3409.6; DB 24; Length 3410;

XX Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAACACAGCTGCAGCGCTGGCTCCGGGTACACAGCGCGCTCGGCCAGGATCTGA 60

DB 1 GGAACACAGCTGCAGCGCTGGCTCCGGGTACACAGCGCGCTCGGCCAGGATCTGA 60

QY 61 GTGATGAGACGTGTCCCACTGAGGTGCCCAACAGCAGAGGTGTGTAGCATGGCTGAG 120

DB 61 GTGATGAGACGTGTCCCACTGAGGTGCCCAACAGCAGAGGTGTGTAGCATGGCTGAG 120

QY 121 AAGCTGGACCGCCACCAAGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGGCAATT 180

DB 121 AAGCTGGACCGCCACCAAGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGGCAATT 180









Db 241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCTGGCCCACTATGTCAGAGGCTGTG 300  
QY 301 GGTGAGCCGCTGCTGGGCAACCGGAAGCCAGCTCTTGTGTTGTAACCTGCTAACCTT 360  
Db 301 GGTGAGCCGCTGCTGGGCAACCGGAAGCCAGCTCTTGTGTTGTAACCTGCTAACCTT 360  
QY 361 TGGCCTGGAGGTGTTTGGCCGAGGCATCACCCTATGTCGCCCTCTGCTGCTGGAAGT 420  
Db 361 TGGCCTGGAGGTGTTTGGCCGAGGCATCACCCTATGTCGCCCTCTGCTGCTGGAAGT 420  
QY 421 GGGGGTAGAGGAGGAGTTATGACCATGTTGCTGGGCATTTGTTCACAGTCTGGGCTGTG 480  
Db 421 GGGGGTAGAGGAGGAGTTATGACCATGTTGCTGGGCATTTGTTCACAGTCTGGGCTGTG 480  
QY 481 CTGTGCTCCGCTCCTTAGGCTACGCCAGTGAACACTGGCCGTGACCGCTATGGCCGCCG 540  
Db 481 CTGTGCTCCGCTCCTTAGGCTACGCCAGTGAACACTGGCCGTGACCGCTATGGCCGCCG 540  
QY 541 GCCCTTCATCTGGGCACTGTCCCTTGGGCATTCCTGTGAGCCCTTCTTCTCATGCCAAGGC 600  
Db 541 GCCCTTCATCTGGGCACTGTCCCTTGGGCATTCCTGTGAGCCCTTCTTCTCATGCCAAGGC 600  
QY 601 CGGCTGGCTAGCAGGCTGCTGTGCCGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
Db 601 CGGCTGGCTAGCAGGCTGCTGTGCCGATCCAGGCCCTGGAGCTGGCACTGCTCAT 660  
QY 661 CTTGGGCTGGGCTGCTGGACTTCTGTGCCAGGTGCTTCACTGACCTGGAGGCCCT 720  
Db 661 CTTGGGCTGGGCTGCTGGACTTCTGTGCCAGGTGCTTCACTGACCTGGAGGCCCT 720  
QY 721 GCTCTCTGACCTTCTCCGGGACCCGACCACTGTGCCAGGCTACTCTGTCTATGCTT 780  
Db 721 GCTCTCTGACCTTCTCCGGGACCCGACCACTGTGCCAGGCTACTCTGTCTATGCTT 780  
QY 781 CATGATCACTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 CATGATCACTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 TGGCCTGGGCTGCTTACCTGGGCAACCGAGGAGTGCCTCTTGGGCTGCTCACTCAT 900  
Db 841 TGGCCTGGGCTGCTTACCTGGGCAACCGAGGAGTGCCTCTTGGGCTGCTCACTCAT 900  
QY 901 CTTCTCTACCTCGGTAGCAGCCACTGCTGTGCTGCTGAGGAGGAGGCTGGGCCCCAC 960  
Db 901 CTTCTCTACCTCGGTAGCAGCCACTGCTGTGCTGCTGAGGAGGAGGCTGGGCCCCAC 960  
QY 961 CGAGCCAGCAGAGGCTGCTGGGCCCTCTGCTGTGCCCCACTGCTGCTGCTGCTGCTGCT 1020  
Db 961 CGAGCCAGCAGAGGCTGCTGGGCCCTCTGCTGTGCCCCACTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 CCGCTTGGCTTTCGGGAACCTGGGCGCTGCTTCCCGGCTGACCAAGCTGCTGCTGCTG 1080  
Db 1021 CCGCTTGGCTTTCGGGAACCTGGGCGCTGCTTCCCGGCTGACCAAGCTGCTGCTGCTG 1080  
QY 1081 CATGCCCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
Db 1081 CATGCCCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 GACCTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Db 1141 GACCTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
QY 1201 AGCTGAGCCGAGCAGGCGCCGAGACACTATGATGAAGCGCTTCGGATGGGCAAGCCT 1260  
Db 1201 AGCTGAGCCGAGCAGGCGCCGAGACACTATGATGAAGCGCTTCGGATGGGCAAGCCT 1260  
QY 1261 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
Db 1261 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 GCAGGATTCGCGCACTCGAGCAGTCTATTGGCCAGTGTGCGAGCTTTCCTGTGGCTGCT 1380  
Db 1321 GCAGGATTCGCGCACTCGAGCAGTCTATTGGCCAGTGTGCGAGCTTTCCTGTGGCTGCT 1380

QY 1381 CGGTGCCACATGCCTGTCCACAGTGTGGCGTGGTGACAGCTTCAGCGCCCTCACCGG 1440  
Db 1381 CGGTGCCACATGCCTGTCCACAGTGTGGCGTGGTGACAGCTTCAGCGCCCTCACCGG 1440  
QY 1441 GTTACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCCTCCCTCTACCAACCGGA 1500  
Db 1441 GTTACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCCTCCCTCTACCAACCGGA 1500  
QY 1501 GAAGCAGGTGTTCCCTGCCCAATACCGAGGGGACACTGGAGTGTAGCAGTGAAGACAG 1560  
Db 1501 GAAGCAGGTGTTCCCTGCCCAATACCGAGGGGACACTGGAGTGTAGCAGTGAAGACAG 1560  
QY 1561 CCTGATGACCAAGCTTCCCTGCCAGCCCTAACCTGGAGCTCCCTTCCCTATGACACAGT 1620  
Db 1561 CCTGATGACCAAGCTTCCCTGCCAGCCCTAACCTGGAGCTCCCTTCCCTATGACACAGT 1620  
QY 1621 GGGTGTGGAGGCTGCTGCCAGCTTCCACCTCCACCGGCTCTGCGGGGCTCTGCGCTG 1680  
Db 1621 GGGTGTGGAGGCTGCTGCCAGCTTCCACCTCCACCGGCTCTGCGGGGCTCTGCGCTG 1680  
QY 1681 TGATGTCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
Db 1681 TGATGTCTCCGTACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
QY 1741 GGGCATCTGCTGACCTCGCCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
Db 1741 GGGCATCTGCTGACCTCGCCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
QY 1801 ATCCCTGTTATGGGCTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
Db 1801 ATCCCTGTTATGGGCTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
QY 1861 TGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
Db 1861 TGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920  
QY 1921 CGACTTGGCCAAATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
Db 1921 CGACTTGGCCAAATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980  
QY 1981 CACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
Db 1981 CACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
QY 2041 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
Db 2041 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
QY 2101 GCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
Db 2101 GCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
QY 2161 ACTGGAGGCTTCCAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
Db 2161 ACTGGAGGCTTCCAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220  
QY 2221 ATGCTGGAATCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Db 2221 ATGCTGGAATCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
QY 2281 CTCTAGTTGAGACACACTAGAGAGGCTTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
Db 2281 CTCTAGTTGAGACACACTAGAGAGGCTTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340  
QY 2341 GTTTCCTATCTTAAGCCCTTAACCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
Db 2341 GTTTCCTATCTTAAGCCCTTAACCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
QY 2401 TTTCTAGGATGAACACTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2460  
Db 2401 TTTCTAGGATGAACACTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 2460

QY 2461 GTCCTGAGGGGCAACACACAGAACACAGGCTCCCTCAGCCACAGCACTGCTTTTGCT 2520  
 DB 2461 GTCCTGAGGGGCAACACACAGAACACAGGCTCCCTCAGCCACAGCACTGCTTTTGCT 2520  
 QY 2521 GATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTGGTCCCTCTGTTGCCATCA 2580  
 DB 2521 GATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTGGTCCCTCTGTTGCCATCA 2580  
 QY 2581 CAGACACAGGCAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640  
 DB 2581 CAGACACAGGCAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640  
 QY 2641 TGCTAGCTTTCTGTGGTGTCTAATATTTGGTAGGGTGGGATCCCAACAATCA 2700  
 DB 2641 TGCTAGCTTTCTGTGGTGTCTAATATTTGGTAGGGTGGGATCCCAACAATCA 2700  
 QY 2701 GGTCCCTGAGATGCTGCTGATGGCTGATCATTGCCAGAATCTTCTTCTCTCGGGT 2760  
 DB 2701 GGTCCCTGAGATGCTGCTGATGGCTGATCATTGCCAGAATCTTCTTCTCTCGGGT 2760  
 QY 2761 GTGGCCCCCAAAATGCTTAACCCAGGACCTTGGAAATTTACTCATCCCAATGATAAT 2820  
 DB 2761 GTGGCCCCCAAAATGCTTAACCCAGGACCTTGGAAATTTACTCATCCCAATGATAAT 2820  
 QY 2821 TCCAAATGCTGTACCCCAAGTTAGGGTGTGAAGGAGGTAGAGGTGGGGCTTCAGGT 2880  
 DB 2821 TCCAAATGCTGTGTACCCCAAGTTAGGGTGTGAAGGAGGTAGAGGTGGGGCTTCAGGT 2880  
 QY 2881 CTCACAGGCTTCCCTTAACCCACCTCTTCTTGTGGCCAGGCTGTGGTCCCTCCCTTCCA 2940  
 DB 2881 CTCACAGGCTTCCCTTAACCCACCTCTTCTTGTGGCCAGGCTGTGGTCCCTCCCTTCCA 2940  
 QY 2941 CTCCTCTACTCTCTAGGACTGGGCTGATGAAGGACCTGCCCAAAATTTCCCTTACC 3000  
 DB 2941 CTCCTCTACTCTCTAGGACTGGGCTGATGAAGGACCTGCCCAAAATTTCCCTTACC 3000  
 QY 3001 CCCAACTTCCCTTACCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 3060  
 DB 3001 CCCAACTTCCCTTACCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 3060  
 QY 3061 GCAGACACAGGACCAAGTGGGTTTCCCAAGCTTGTCCATCTCAGCCCTCAGCTAGG 3120  
 DB 3061 GCAGACACAGGACCAAGTGGGTTTCCCAAGCTTGTCCATCTCAGCCCTCAGCTAGG 3120  
 QY 3121 ATATCTGTCTTGGGAAATCTCACACAGAACTCAGGACACCTTGTCCATCTCAGCCCTCAGCTAGG 3180  
 DB 3121 ATATCTGTCTTGGGAAATCTCACACAGAACTCAGGACACCTTGTCCATCTCAGCCCTCAGCTAGG 3180  
 QY 3181 GAGGTCTTATCTCTCAGGGGGGTTTAAAGTGGCGTTTGAATATGCTGCTTATTATT 3240  
 DB 3181 GAGGTCTTATCTCTCAGGGGGGTTTAAAGTGGCGTTTGAATATGCTGCTTATTATT 3240  
 QY 3241 TAGCGGGTGAATATTTTATCTGTAAGTGAGCAATCAGATATATGTTTATGTTGACA 3300  
 DB 3241 TAGCGGGTGAATATTTTATCTGTAAGTGAGCAATCAGATATATGTTTATGTTGACA 3300  
 QY 3301 AATTAAGGCTTCTTATATGTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3360  
 DB 3301 AATTAAGGCTTCTTATATGTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3360  
 QY 3361 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3410  
 DB 3361 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3410

RESULT 11  
 ID AAA06349  
 XX AAA06349 standard; cdna; 3410 BP.  
 AC AAA06349;  
 XX  
 DT 13-JUN-2000 (first entry)  
 XX

Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:110.  
 Human; prostate cancer; diagnosis; tumour; gene therapy; detection;  
 Immunogenic; cytostatic; vaccine; ss.  
 Homo sapiens.  
 WO200004149-A2.  
 XX 27-JAN-2000.  
 XX 14-JUL-1999; 99WO-US15838.  
 XX 14-JUL-1998; 98US-0115453.  
 PR 14-JUL-1998; 98US-0116134.  
 PR 14-SEP-1998; 98US-0159812.  
 PR 23-SEP-1998; 98US-0159822.  
 PR 15-JAN-1999; 99US-0232149.  
 PR 15-JAN-1999; 99US-0232880.  
 PR 09-APR-1999; 99US-0288946.  
 XX (CORI-) CORIXA CORP.  
 PA Dillon DC, Harlocker SL, Yuqiu J, Xu J, Mitcham JL;  
 PI WPI; 2000-171268/15.  
 XX New polypeptide useful for treating and diagnosing prostate cancer  
 PT comprises an immunogenic portion of prostate tumor protein -  
 PS Claim 1; Page 135-136; 263pp; English.  
 XX The present invention describes isolated polypeptides, comprising an  
 CC immunogenic portion of a prostate tumour protein (pmp). The polypeptides  
 CC and polynucleotides encoding them have cytostatic activity and can be  
 CC used in vaccines and in gene therapy. The polypeptides and  
 CC polynucleotides encoding them, antigen presenting cells which express  
 CC the polypeptides, antibodies against the polypeptides and vaccines  
 CC comprising them can be used for inhibiting the development of prostate  
 CC cancer in a patient. The polypeptides can be used to generate antibodies  
 CC or anti-idiotypic antibodies for passive immuno therapy. A portion of  
 CC the polynucleotides encoding the polypeptides can be used as a probe or  
 CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and  
 CC AAY82000 to AAY82020 represent sequences used in the exemplification of  
 CC the present invention.  
 XX Sequence 3410 BP; 667 A; 1015 C; 945 G; 782 T; 1 other;  
 SQ  
 Query Match 99.9%; Score 3408; DB 21; Length 3410;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3409; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GGGAAACAGCTGACAGCGCTGGCTCCGGGTGACAGCGCGCGCTGGCCAGGATCTGA 60  
 DB 1 GGGAAACAGCTGACAGCGCTGGCTCCGGGTGACAGCGCGCGCTGGCCAGGATCTGA 60  
 QY 61 GTGATGAGAGCTGCTCCCACTGAGTGCCCAACAGCAGCAGGTGTGAGCATGGGCTGAG 120  
 DB 61 GTGATGAGAGCTGCTCCCACTGAGTGCCCAACAGCAGCAGGTGTGAGCATGGGCTGAG 120  
 QY 121 AAGTGGACAGGAGGAGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGGCACTT 180  
 DB 121 AAGTGGACAGGAGGAGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGGCACTT 180  
 QY 181 GCGGACAGGAGGAGGAGGCTGGCAGCTTCTGGAGCAGAGCCGAGAGCAAGAGCACTTCTG 240  
 DB 181 GCGGACAGGAGGAGGAGGCTGGCAGCTTCTGGAGCAGAGCCGAGAGCAAGAGCACTTCTG 240  
 QY 241 GAGTGCCTGAGCGGCGCTTACCGGCTTACCGGCTTACCGGCTTACCGGCTTACCGGCTTAC 300  
 DB 241 GAGTGCCTGAGCGGCGCTTACCGGCTTACCGGCTTACCGGCTTACCGGCTTACCGGCTTAC 300  
 QY 301 GGTGAGCGGCTGCTGGGCGCAGCCGAGAAAGCCAGCTCTTGTGCTCAACCTGCTTAACCTT 360

Db 301 GGTGAGCCGGCTGCTGGCGACCGAAGCCAGCTCTTGGTGAACCTGCTAACCTT 360  
Qy 361 TGGCCTGGAGGTGTGTTGGCGCAGGACATCACTATATGCGCCCTCTGCTGCTGAACT 420  
Db 361 TGGCCTGGAGGTGTGTTGGCGCAGGACATCACTATGTCGCCCTCTGCTGCTGAACT 420  
Qy 421 GGGGGTAGAGGAGAAAGTTACATGACATGTGTGCTGGCATTGTPCCAGTGTGGCCCTGTG 480  
Db 421 GGGGGTAGAGGAGAAAGTTACATGACATGTGTGCTGGCATTGTPCCAGTGTGGCCCTGTG 480  
Qy 481 CTGTGTCCCGCTCTAGGCTACGCGCTACGCTAGGCTGGCGTGTGACGCTATGCGCGCGCGG 540  
Db 481 CTGTGTCCCGCTCTAGGCTACGCGCTACGCTAGGCTGGCGTGTGACGCTATGCGCGCGCGG 540  
Qy 541 GCCCTTCATCTGGGCACTGTCTTGGGCATCTCTGCTGAGCTCTTCTCATCCCAAGGCG 600  
Db 541 GCCCTTCATCTGGGCACTGTCTTGGGCATCTCTGCTGAGCTCTTCTCATCCCAAGGCG 600  
Qy 601 CGGCTGGCTAGCAGGCTGTGTGCGCGGATCCAGGCCCCCTGGAGCTGGCACTGCTCAT 660  
Db 601 CGGCTGGCTAGCAGGCTGTGTGCGCGGATCCAGGCCCCCTGGAGCTGGCACTGCTCAT 660  
Qy 661 CTTGGGCTGGGCTGTCTGAGCTTCTGTGGCCAGGTGTCTTCACTCCACTGAGGCTCT 720  
Db 661 CTTGGGCTGGGCTGTCTGAGCTTCTGTGGCCAGGTGTCTTCACTCCACTGAGGCTCT 720  
Qy 721 GCTCTCTGACCTTCTCCGGGACCCGACCACTGTGTGCGCAGGCTACTCTGTCTATGCTT 780  
Db 721 GCTCTCTGACCTTCTCCGGGACCCGACCACTGTGTGCGCAGGCTACTCTGTCTATGCTT 780  
Qy 781 CATGATCAGTCTTGGGGCTGTCTGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 CATGATCAGTCTTGGGGCTGTCTGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Qy 841 TGCCCTGGCCCTACTCTGGGACCCAGGAGGAGTCTCTTGGCTGCTGCTGCTGCTGCT 900  
Db 841 TGCCCTGGCCCTACTCTGGGACCCAGGAGGAGTCTCTTGGCTGCTGCTGCTGCTGCT 900  
Qy 901 CTCTCTCACTGGTAGCAGCAGCAGTCTGTGTGCTGCTGAGGAGGAGGCTGCGGCGCAC 960  
Db 901 CTCTCTCACTGGTAGCAGCAGCAGTCTGTGTGCTGCTGAGGAGGAGGCTGCGGCGCAC 960  
Qy 961 CGAGCCAGCAGAGGCTGTGCGCCCTCTCTGCTGCGCCCACTGCTGCTCATGCGGCG 1020  
Db 961 CGAGCCAGCAGAGGCTGTGCGCCCTCTCTGCTGCGCCCACTGCTGCTCATGCGGCG 1020  
Qy 1021 CGGCTTGGCTTTCGGGAACCTGGGCGCCCTGCTTCCCGGCTGCACTGCTGCTGCTG 1080  
Db 1021 CGGCTTGGCTTTCGGGAACCTGGGCGCCCTGCTTCCCGGCTGCACTGCTGCTGCTG 1080  
Qy 1081 CATGCCCCCAGCTGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 1081 CATGCCCCCAGCTGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Qy 1141 GACCTTACGCTGTTTACAGGATTTCTGGGAGGAGGCTGTACAGGCGCTGCGCCAG 1200  
Db 1141 GACCTTACGCTGTTTACAGGATTTCTGGGAGGAGGCTGTACAGGCGCTGCGCCAG 1200  
Qy 1201 AGCTGAGCGGCGACCGAGGCGCGGAGACATATGATGAAGGCTTGGATGGGCGAGCT 1260  
Db 1201 AGCTGAGCGGCGACCGAGGCGCGGAGACATATGATGAAGGCTTGGATGGGCGAGCT 1260  
Qy 1261 GGGGCTGTCTGCACTGCGCATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
Db 1261 GGGGCTGTCTGCACTGCGCATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
Qy 1321 GCAGCGANTTCGCACTCGAGAGCTATTTGGCGAGTGTGGCAGCTTTCCTGTGGCTG 1380  
Db 1321 GCAGCGANTTCGCACTCGAGAGCTATTTGGCGAGTGTGGCAGCTTTCCTGTGGCTG 1380  
Qy 1381 GGTGCGACATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
Db 1381 GGTGCGACATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

Db 1381 CGGTGCCACATGCTGCTCCACAGTGTGGCGTGTGACAGCTTACGCGCGCTCACCGG 1440  
Qy 1441 GTTACCTTCTCAGCCCTGCGATCCTGCCCTTACACTAGGCTTCCCTTCTACCAACCGGGA 1500  
Db 1441 GTTACCTTCTCAGCCCTGCGATCCTGCCCTTACACTAGGCTTCCCTTCTACCAACCGGGA 1500  
Qy 1501 GAAGCAGGTGTCTGCGCCAAATACGAGGAGGACACTGAGGCTGCTAGCAGTGAAGACAG 1560  
Db 1501 GAAGCAGGTGTCTGCGCCAAATACGAGGAGGACACTGAGGCTGCTAGCAGTGAAGACAG 1560  
Qy 1561 CTTGATGACCACTTCTGCGAGGCTTAAGCCTGAGCTCCCTTCCCTAAATGACACAGT 1620  
Db 1561 CTTGATGACCACTTCTGCGAGGCTTAAGCCTGAGCTCCCTTCCCTAAATGACACAGT 1620  
Qy 1621 GGTGCTGGAGGCTGCTGCTGCCACCTTCCACCGCGCTCTGCGGCGCTCTGCGCTG 1680  
Db 1621 GGTGCTGGAGGCTGCTGCTGCCACCTTCCACCGCGCTCTGCGGCGCTCTGCGCTG 1680  
Qy 1681 TGATGCTTCCCTAGTGTGTGGTGAAGCCACAGGAGGCTGCTGCGGCGCTCTGCGCTG 1740  
Db 1681 TGATGCTTCCCTAGTGTGTGGTGAAGCCACAGGAGGCTGCTGCGGCGCTCTGCGCTG 1740  
Qy 1741 GGGCATCTGCTGACCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
Db 1741 GGGCATCTGCTGACCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
Qy 1801 ATCCCTGTTTATGGCTCCATTTGCTCAGCTCAGCAGTCTGCTGCTGCTGCTGCTGCTGCT 1860  
Db 1801 ATCCCTGTTTATGGCTCCATTTGCTCAGCTCAGCAGTCTGCTGCTGCTGCTGCTGCTGCT 1860  
Qy 1861 TGCGCGAGGCTGGTGTGGTGAAGTACTTACTTGTGTACACAGTACTATTTGACAAGAG 1920  
Db 1861 TGCGCGAGGCTGGTGTGGTGAAGTACTTACTTGTGTACACAGTACTATTTGACAAGAG 1920  
Qy 1921 CGACTTGGCCAAATACCTCAGCTAGAACTTCCAGCACATTTGGGCTGAGGCGCTGCT 1980  
Db 1921 CGACTTGGCCAAATACCTCAGCTAGAACTTCCAGCACATTTGGGCTGAGGCGCTGCT 1980  
Qy 1981 CACTGGTCCAGCTCCCGCTCTCTGTTAGCCCATGGGCTGCGGCGCTGCGGCGCTGCT 2040  
Db 1981 CACTGGTCCAGCTCCCGCTCTCTGTTAGCCCATGGGCTGCGGCGCTGCGGCGCTGCT 2040  
Qy 2041 TTTCTGTTGCTGCCAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
Db 2041 TTTCTGTTGCTGCCAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
Qy 2101 GCTGACAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
Db 2101 GCTGACAGCTGGGGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
Qy 2161 ACTGGAGGCTTCCAGGGGCTTCTAGTGTGAGTATACAGGAGGCGCCAGAGGCTGCT 2220  
Db 2161 ACTGGAGGCTTCCAGGGGCTTCTAGTGTGAGTATACAGGAGGCGCCAGAGGCTGCT 2220  
Qy 2221 ATGCACTTGAATCGGGGCTGCTGAGTGTGATTTACAGGCTCAGGCTTACAGCTAGC 2280  
Db 2221 ATGCACTTGAATCGGGGCTGCTGAGTGTGATTTACAGGCTCAGGCTTACAGCTAGC 2280  
Qy 2281 CTCTAGTTGAGACACACTTAGAGAGGTTTTTGGAGCTGAATAAAGCTCAGTCACTG 2340  
Db 2281 CTCTAGTTGAGACACACTTAGAGAGGTTTTTGGAGCTGAATAAAGCTCAGTCACTG 2340  
Qy 2341 GTTTCCTCTTAAGCCCTTAAGCTGAGCTTCTGTTTAAAGTACTCTTGCATGGGAG 2400  
Db 2341 GTTTCCTCTTAAGCCCTTAAGCTGAGCTTCTGTTTAAAGTACTCTTGCATGGGAG 2400  
Qy 2401 TTTCTAGGATGAACACTCTCTCCATGGGATTTGAACATATGACTTATTTGTAGGGAGA 2460  
Db 2401 TTTCTAGGATGAACACTCTCTCCATGGGATTTGAACATATGACTTATTTGTAGGGAGA 2460  
Qy 2461 GTCTGAGGGGCAACACAGAACAGGTCCTGCGCTCAGCCCCACAGCACTGTCTTTTGT 2520  
Db 2461 GTCTGAGGGGCAACACAGAACAGGTCCTGCGCTCAGCCCCACAGCACTGTCTTTTGT 2520

QY	2521	GATCACCC	CCCTCTT	TA	CTTTTATCAGGATG	GC	CGCTGTGGT	CCTCTCTGTTGCCATCA	2580
Db	2521	GATCACCC	CCCTCTT	TA	CTTTTATCAGGATG	GC	CGCTGTGGT	CCTCTCTGTTGCCATCA	2580
QY	2581	CAGACACAGC	GCA	TTTAAA	TATTTAACTTATTTAA	CAAAAGT	TAG	AAAGGAATCCAT	2640
Db	2581	CAGACACAGC	GCA	TTTAAA	TATTTAACTTATTTAA	CAAAAGT	TAG	AAAGGAATCCAT	2640
QY	2641	TGCTAGCTTT	CTCTGT	TGGTGTCTAA	TATTTTGGGTAG	GGTGGGGAT	CCCCAACAA	TCA	2700
Db	2641	TGCTAGCTTT	CTCTGT	TGGTGTCTAA	TATTTTGGGTAG	GGTGGGGAT	CCCCAACAA	TCA	2700
QY	2701	GGTCCCTGAG	ATAGCT	GGTCTA	TNTGGGCTGATCA	TG	CCAGAA	TCTTCTTCC	2760
Db	2701	GGTCCCTGAG	ATAGCT	GGTCTA	TNTGGGCTGATCA	TG	CCAGAA	TCTTCTTCC	2760
QY	2761	CTGGCCCC	CCCAAA	TGCCTAA	CCAGGACCTTGG	AAATTTCTACTCAT	CCCCAAATGATA	T	2820
Db	2761	CTGGCCCC	CCCAAA	TGCCTAA	CCAGGACCTTGG	AAATTTCTACTCAT	CCCCAAATGATA	T	2820
QY	2821	TCCAAATGCT	TTACCC	AAAGTTAG	GGTGTTCGA	AGGAAGGTAG	AGGGTGGGGCTTC	CAGGT	2880
Db	2821	TCCAAATGCT	TTACCC	AAAGTTAG	GGTGTTCGA	AGGAAGGTAG	AGGGTGGGGCTTC	CAGGT	2880
QY	2881	CTCAAGGGCT	TCCTTA	ACAC	CCCTCTTCTTTGG	CCAGCCTGGT	TCC	CCACCTTCCA	2940
Db	2881	CTCAAGGGCT	TCCTTA	ACAC	CCCTCTTCTTTGG	CCAGCCTGGT	TCC	CCACCTTCCA	2940
QY	2941	CTCCCTCT	ACTCTCT	CTAGGACT	GGGCTGATGA	AGGCAC	TGCC	CCAAAATTTCC	3000
Db	2941	CTCCCTCT	ACTCTCT	CTAGGACT	GGGCTGATGA	AGGCAC	TGCC	CCAAAATTTCC	3000
QY	3001	CCCACTTT	CCCTAC	CC	CCCACTTTCC	CCAC	CGCTCC	CAACCCCTGTTG	3060
Db	3001	CCCACTTT	CCCTAC	CC	CCCACTTTCC	CCAC	CGCTCC	CAACCCCTGTTG	3060
QY	3061	GCAGGACCA	GACAGCA	AAAGTGG	GGTTTCC	CAAGCCTTTG	TCCATCTC	AGCCCCCAGAGT	3120
Db	3061	GCAGGACCA	GACAGCA	AAAGTGG	GGTTTCC	CAAGCCTTTG	TCCATCTC	AGCCCCCAGAGT	3120
QY	3121	ATATCTGT	CTTGGG	AAATCT	CACACAGAA	ACTCAGGAGC	ACCCCTGCTG	CAGCTAAGG	3180
Db	3121	ATATCTGT	CTTGGG	AAATCT	CACACAGAA	ACTCAGGAGC	ACCCCTGCTG	CAGCTAAGG	3180
QY	3181	GAGTCTTAT	CTCTCAG	GGGGGTTTTA	AGTTCGGGTTT	GCAATTAAT	TCGCTTATTTAT	T	3240
Db	3181	GAGTCTTAT	CTCTCAG	GGGGGTTTTA	AGTTCGGGTTT	GCAATTAAT	TCGCTTATTTAT	T	3240
QY	3241	TAGCGGGGT	GAA	TATTTTAT	CTGTAAGT	GAGCAATCAG	AGTATATGTT	TATGGTGACA	3300
Db	3241	TAGCGGGGT	GAA	TATTTTAT	CTGTAAGT	GAGCAATCAG	AGTATATGTT	TATGGTGACA	3300
QY	3301	AAATTAAG	CGCTTCTT	TATATGTTT	TAAAA	AAAAA	AAAAA	AAAAA	

RESULT 12  
AAS14962  
ID AAS14962 standard; cDNA; 3320 BP.  
XX  
XX AAS14962;  
XX  
XX 14-FEB-2002 (first entry)  
DT  
XX  
XX DNA encoding human PROST 03.  
DE  
XX  
XX Human; PROST 03; metastasis; prostate cancer; immune response;  
KW tumour;

KW	cytostatic; vaccine; ss.
XX	
OS	Homo sapiens.
XX	
XX	Location/Qualifiers
PH	282..1943
FT	/*tag= a
FT	/product= "PROST 03"
FT	
XX	
XX	WO200181577-A2.
XX	
XX	
PD	01-NOV-2001.
XX	
PF	26-APR-2001; 2001WO-US13323.
XX	
XX	27-APR-2000; 2000US-200065P.
PR	20-APR-2001; 2001US-0200065.
PR	
XX	(SCHD ) SCHERING AG.
PA	
XX	
PI	Lau T, Lin RJ, Parkes D, Parry G, Schneider DW, Steinbrecher R;
PI	Van Héult PT, Wu J;
XX	
DR	WPI; 2002-041404/05.
DR	P-PSDB; AAU10324.
XX	
PT	Novel PROST 03 polypeptides and polynucleotides useful in research,
PT	diagnosis and therapeutic applications, particularly for use in cancer
PT	therapeutics
XX	
PS	Claim 6; Fig 1; 77pp; English.
XX	
CC	The invention relates to an isolated PROST 03 polypeptide (I) and to
CC	the polynucleotide (II) encoding PROST 03. Fragments of (I) were used
CC	to generate antibodies (III) to PROST 03. (III) is useful for selectively
CC	destroying a cell expressing (I), and for treating a disease-state
CC	associated with expression of PROST 03 in a human patient. (III) is
CC	useful for diagnosing metastasis associated with (I), in a subject.
CC	(I) is also useful for diagnosing and treating diseases of cell
CC	proliferation such as prostate cancers. (I) is also useful for generating
CC	antibodies to PROST 03. (III) is useful in detecting the levels of PROST.
CC	03 polypeptides in cells and tissues, and in targeting drugs to primary
CC	and metastatic tumours. (I) is also useful for stimulating immune
CC	response to PROST 03 containing cells. (II) is useful in diagnostic
CC	assays for detecting the levels of polynucleotides encoding PROST 03 in
CC	cells and tissues. (II) is useful as DNA probes, as targets for antisense
CC	and ribozyme therapy, and as templates for the production of antisense
CC	polynucleotides. (I) and (II) are useful in research, biological,
CC	clinical and therapeutic purposes. The present sequence represents
CC	the coding sequence of human PROST 03.

Sequence 3320 BP: 585 A: 1013 C: 944 G: 778 T: 0 other: XX

	Query Match	96.6%	Score 3292.4;	DB 24;	Length 3320;	
	Best Local Similarity	99.9%;	Pred. No. 0;			
	Matches 3316;	Conservative 0;	Mismatches 1;	Indels 3;	Gaps 2;	
Qy	3	GAACCAAGCTCAGCGCTGCTCCGGGTGACAGCGCGCGCCCTCGGCCAGATCTGAGT	62			
Dd	1	GAACCAAGCTCAGCGCTGCTCCGGGTGACAGCGCGCGCCCTCGGCCAGATCTGAGT	60			
Qy	63	GATGAGAGTGTCCTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGATGGCGTGAGAA	122			
Dd	61	GATGAGAGTGTCCTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGCGTGAGAA	120			
Qy	123	GCTGGACCGGCACCAGAAGGGCTGGCAGAAATGGCGCCTGCTGATTCTTAGGCAGTTGG	182			
Dd	121	GCTGGACCGGCACCAGAAGGGCTGGCAGAAATGGCGCCTGCTGATTCTTAGGCAGTTGG	180			
Qy	183	CGGCAGCAAGGAGGAGGGCCGACGTTCTGGAGCAGAGCCGACGAAGCAGTTCTCGGA	242			
Dd	181	CGGCAGCAAGGAGGAGGGCCGACGTTCTGGAGCAGAGCCGACGAAGCAGTTCTCGGA	240			

QY 243 GTGCGCTGAACGGCCCTGAGCCCTACCGCGCTGGCCCACTATGCTCCAGAGGCTGTGGG 302  
DB 241 GTGCGCTGAACGGCCCTGAGCCCTACCGCGCTGGCCCACTATGCTCCAGAGGCTGTGGG 300  
QY 303 TGAGCGCGCTGCTGGGGCACCAGAAAGCCAGCTCTTGCTGCTCAACCTGCTAACTTTG 362  
DB 301 TGAGCGCGCTGCTGGGGCACCAGAAAGCCAGCTCTTGCTGCTCAACCTGCTAACTTTG 360  
QY 363 GCGTGAAGTGTGTTGGCGCGCAGGAGCATCACTATGTGCGCGCTCTGCTGCTGAAGTGG 422  
DB 361 GCGTGAAGTGTGTTGGCGCGCAGGAGCATCACTATGTGCGCGCTCTGCTGCTGAAGTGG 420  
QY 423 GGTGTAGAGAGAAGTTTCATGACCATGCTGGGCAATGGTCCAGTGTGGCGCTGGTCT 482  
DB 421 GGTGTAGAGAGAAGTTTCATGACCATGCTGGGCAATGGTCCAGTGTGGCGCTGGTCT 480  
QY 483 GTGTCCCGCTCCTAGGCTCAGCAGTACCACTGGCGTGGAGCTATGGCGCGCGCGCGC 542  
DB 481 GTGTCCCGCTCCTAGGCTCAGCAGTACCACTGGCGTGGAGCTATGGCGCGCGCGCGC 540  
QY 543 CCTTCATCTGGGCACTGCTTGGGCATCTGCTGAGCCTCTTCTCATGCCAAGGGCGG 602  
DB 541 CCTTCATCTGGGCACTGCTTGGGCATCTGCTGAGCCTCTTCTCATGCCAAGGGCGG 600  
QY 603 GCTGCTAGCAGGGCTGCTGTGCCCGGATCCAGGCCCTGGAGCTGSCACTGCTCATCC 662  
DB 601 GCTGCTAGCAGGGCTGCTGTGCCCGGATCCAGGCCCTGGAGCTGSCACTGCTCATCC 660  
QY 663 TGGGGTGGGGTGTGCTGACCTTCTGTGCCAGGTGTGCTTCACTCCACTGAGGCCCTTC 722  
DB 661 TGGGGTGGGGTGTGCTGACCTTCTGTGCCAGGTGTGCTTCACTCCACTGAGGCCCTTC 720  
QY 723 TCTCTGACCTCTCCGGGACCGGACCACTGTCCGAGGCTACTCTCTATGCCCTCA 782  
DB 721 TCTCTGACCTCTCTCCGGGACCGGACCACTGTCCGAGGCTACTCTCTATGCCCTCA 780  
QY 783 TGATCAGCTTTGGGGCTGCTGGGCTACCTCTGCTGCCATTTGACTGGGACACCACTG 842  
DB 781 TGATCAGCTTTGGGGCTGCTGGGCTACCTCTGCTGCCATTTGACTGGGACACCACTG 840  
QY 843 CCTTGGCCCCCTACCTGGGACCCAGGAGGAGTGGCTTTGGCGCTGCTACCCCTCATCT 902  
DB 841 CCTTGGCCCCCTACCTGGGACCCAGGAGGAGTGGCTTTGGCGCTGCTACCCCTCATCT 900  
QY 903 TCTTCACCTGCTAGCAGCACACTGCTGGTGGTGGAGGAGCAGCGCTGGGCCACCG 962  
DB 901 TCTTCACCTGCTAGCAGCACACTGCTGGTGGTGGAGGAGCAGCGCTGGGCCACCG 960  
QY 963 AGCCAGCAGAAAGGCTGTGCGGCCGCCCTCTTGTGCGCCCCACTGCTGCTGCGGGCCC 1022  
DB 961 AGCCAGCAGAAAGGCTGTGCGGCCGCCCTCTTGTGCGCCCCACTGCTGCTGCGGGCCC 1020  
QY 1023 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCGGCTGCACAGTGTGCTGCCGCA 1082  
DB 1021 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCGGCTGCACAGTGTGCTGCCGCA 1080  
QY 1083 TGCCCGCCGACCTGCGCGGCTCTGCTGGGTGAGCTGTGAGCTGGATGSCACTCATGA 1142  
DB 1081 TGCCCGCCGACCTGCGCGGCTCTGCTGGGTGAGCTGTGAGCTGGATGSCACTCATGA 1140  
QY 1143 CTTTCACGCTGTTTACACGATTTCTGCTGGCGAGGGCTGTATACAGGGGTGCGCCAGAG 1202  
DB 1141 CTTTCACGCTGTTTACACGATTTCTGCTGGCGAGGGCTGTATACAGGGGTGCGCCAGAG 1200  
QY 1203 CTGAGCCGGGACCCAGGCGCGGAGACACTATGATGAAGGGCTTCGATGGGACGCTGG 1262  
DB 1201 CTGAGCCGGGACCCAGGCGCGGAGACACTATGATGAAGGGCTTCGATGGGACGCTGG 1260  
QY 1263 GCGTCTTCTGAGTGGCCATCTCCCTGCTGCTTCTCTGCTGATGACGCGCTGGTGC 1322  
DB 1261 GCGTCTTCTGAGTGGCCATCTCCCTGCTGCTTCTCTGCTGATGACGCGCTGGTGC 1320  
QY 1323 AGCGATTGGGCACTCGAGCAGTCTATTGTCGCGAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1382

DB 1321 AGCGATTGGGCACTCGAGCAGTCTATTGTCGCGAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1380  
QY 1383 GTGCCACATGCTGCTGCCACAGTGTGGCGCTGTGACAGCTTTCAGCGCGCTCACCAGGT 1442  
DB 1381 GTGCCACATGCTGCTGCCACAGTGTGGCGCTGTGACAGCTTTCAGCGCGCTCACCAGGT 1440  
QY 1443 TCACCTTCTCAGCCCTCAGATCCTGCGCTTACACACTTGGCCTCCCTCTACACCGGGAGA 1502  
DB 1441 TCACCTTCTCAGCCCTCAGATCCTGCGCTTACACACTTGGCCTCCCTCTACACCGGGAGA 1500  
QY 1503 AGCAGGTGTTCTGCGCCCAATACGAGGGGACACTGGAGGTGCTAGCAGTGAAGACAGCC 1562  
DB 1501 AGCAGGTGTTCTGCGCCCAATACGAGGGGACACTGGAGGTGCTAGCAGTGAAGACAGCC 1560  
QY 1563 TGATGACACAGTTCCTGCCAGGCGCTAAGCTCTGAGCTTCCCTTCCCTAATGACACGTGG 1622  
DB 1561 TGATGACACAGTTCCTGCCAGGCGCTAAGCTCTGAGCTTCCCTTCCCTAATGACACGTGG 1620  
QY 1623 GTGCTGAGGACGTGGCCTGCTCCCACTCCACCGCGCTCTCGGGGCGCTCTGCCCTGTG 1682  
DB 1621 GTGCTGAGGACGTGGCCTGCTCCCACTCCACCGCGCTCTCGGGGCGCTCTGCCCTGTG 1680  
QY 1683 ATGTCTCCGTACGTGGTGGTGGTGACCCACCGAGGCGAGGCTGGTTCGCGGCGCGG 1742  
DB 1681 ATGTCTCCGTACGTGGTGGTGGTGACCCACCGAGGCGAGGCTGGTTCGCGGCGCGG 1740  
QY 1743 GCATCTGCTGAGCCTCGCCATCTGATAGTGCCTTCTGCTGCTCCAGGTGGCCCAT 1802  
DB 1741 GCATCTGCTGAGCCTCGCCATCTGATAGTGCCTTCTGCTGCTCCAGGTGGCCCAT 1800  
QY 1803 CCCTGTTATGGGCTCCATGTCCAGCTCAGCCAGTCTGTCACTGCTATATGGTGTCTG 1862  
DB 1801 CCCTGTTATGGGCTCCATGTCCAGCTCAGCCAGTCTGTCACTGCTATATGGTGTCTG 1860  
QY 1863 CCGCAGGCTGGGCTGCTGGTGGCGCATTTACTTTCCTACACAGTGTGTTGACACAGCG 1922  
DB 1861 CCGCAGGCTGGGCTGCTGGTGGCGCATTTACTTTCCTACACAGTGTGTTGACACAGCG 1920  
QY 1923 ACTTGGCCAAATACTCAGCGTAGAAACTTCCAGCACATTTGGGGTGGAGGCGCTGCTCA 1982  
DB 1921 ACTTGGCCAAATACTCAGCGTAGAAACTTCCAGCACATTTGGGGTGGAGGCGCTGCTCA 1980  
QY 1983 CTGGTCCAGCTCCCGCTCCTGTAGCCCATGGGGCTGCGCGCTGCGCGCCAGTTT 2042  
DB 1981 CTGGTCCAGCTCCCGCTCCTGTAGCCCATGGGGCTGCGCGCTGCGCGCCAGTTT 2040  
QY 2043 CTGTTGCTGCAAGTAATGTGGCTCTGCTGCCACCTGTGCTGCTGAGGTGCGTAGC 2102  
DB 2041 CTGTTGCTGCAAGTAATGTGGCTCTGCTGCCACCTGTGCTGCTGAGGTGCGTAGC 2100  
QY 2103 TGACAGCTGGGGGCTGGGGCTGCTCCTCTCTCTCCCCAGTCTCTAGGGTGGCTGAC 2162  
DB 2101 TGACAGCTGGGGGCTGGGGCTGCTCCTCTCTCTCCCCAGTCTCTAGGGTGGCTGAC 2160  
QY 2163 TGAGGCGCTTCCAAAGGGGTTTCAGTCTGACTTATACAGGAGGCGCAAGGGCTCCAT 2222  
DB 2161 TGAGGCGCTTCCAAAGGGGTTTCAGTCTGACTTATACAGGAGGCGCAAGGGCTCCAT 2220  
QY 2223 GCATGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGCTTAACAGCTAGCCT 2282  
DB 2221 GCATGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGCTTAACAGCTAGCCT 2280  
QY 2283 CTTAGTTGAGACACACTAGAGAAGGGTTTTTGGGAGCTGGAATAAAGTCACTGCTGCT 2342  
DB 2281 CTTAGTTGAGACACACTAGAGAAGGGTTTTTGGGAGCTGGAATAAAGTCACTGCTGCT 2340  
QY 2343 TTCCCATCTCTAGGCCCTTAACCTGCAGCTTCGTTTAAATGAGTCTTTCATGAGGAGTT 2402  
DB 2341 TTCCCATCTCTAGGCCCTTAACCTGCAGCTTCGTTTAAATGAGTCTTTCATGAGGAGTT 2400  
QY 2403 TCTAGATGAACACACTCTCCCTCCATGGGATTTGAACATATG--ACTTATTTTGTAGGGAGA 2460  
DB 2401 TCTAGATGAACACACTCTCCCTCCATGGGATTTGAACATATG--ACTTATTTTGTAGGGAGA 2460



Db 2401 TCTAGATGAACACCTCTCCATGGGATTTGAACATATGAAGTTATTGTAGGGGAAGA 2460  
 QY 2461 GTCTGAGGGCAACACACAGAACAGGTCCCTCAGCCACACAGCTGTCTTTTGGCT 2520  
 Db 2461 GTCTGAGGGCAACACACAGAACAGGTCCCTCAGCCACACAGCTGTCTTTTGGCT 2520  
 QY 2521 GATCCACCCCTCTTACCTTTATCAGATGTGGCTGTGGTCTCTGTGTGCAATCA 2580  
 Db 2521 GATCCACCCCTCTTACCTTTATCAGATGTGGCTGTGGTCTCTGTGTGCAATCA 2580  
 QY 2581 CAGAGACACAGCAATTAATATTAACCTATTATTAACAAAGTAGAAGGAATCCAT 2640  
 Db 2581 CAGAGACACAGCAATTAATATTAACCTATTATTAACAAAGTAGAAGGAATCCAT 2640  
 QY 2641 TGCTAGCTTTCTGTGTTGCTCTAAATATTTGGGTAGGGTGGGGATCCCAACAATCA 2700  
 Db 2641 TGCTAGCTTTCTGTGTTGCTCTAAATATTTGGGTAGGGTGGGGATCCCAACAATCA 2700  
 QY 2701 GTTCCCTCAGATAGTGTCTATTTGGCTGATCATTTGCCAGAAATCTTCTCTCTGGGT 2760  
 Db 2701 GTTCCCTCAGATAGTGTCTATTTGGCTGATCATTTGCCAGAAATCTTCTCTCTGGGT 2760  
 QY 2761 CTGGCCCTTACTCTCTAGGACTGGCTGATGAAGGCACTGCCCAAAATTTCCCTACC 2820  
 Db 2761 CTGGCCCTTACTCTCTAGGACTGGCTGATGAAGGCACTGCCCAAAATTTCCCTACC 2820  
 QY 2821 TCCAAATGCTTTACCCAAAGTTAGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880  
 Db 2821 TCCAAATGCTTTACCCAAAGTTAGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880  
 QY 2881 CTCAGGGTTCCTTAACACCCCTCTCTCTTTGGCCAGCTGTTTCCCTCCACTTCCA 2940  
 Db 2881 CTCAGGGTTCCTTAACACCCCTCTCTCTTTGGCCAGCTGTTTCCCTCCACTTCCA 2940  
 QY 2941 CTCCCTTACTCTCTAGGACTGGCTGATGAAGGCACTGCCCAAAATTTCCCTACC 3000  
 Db 2941 CTCCCTTACTCTCTAGGACTGGCTGATGAAGGCACTGCCCAAAATTTCCCTACC 3000  
 QY 3001 CCCAACTTTCCCTACCCCAACTTTCCCAAGCTCCACAACTCTTTGGAGCTACT 3060  
 Db 3001 CCCAACTTTCCCTACCCCAACTTTCCCAAGCTCCACAACTCTTTGGAGCTACT 3060  
 QY 3061 GCAGGACAGACAGCAAGTGGGTTTCCCAAGCTTTTGTCCATCTCAGCCCCCAGAGT 3120  
 Db 3061 GCAGGACAGACAGCAAGTGGGTTTCCCAAGCTTTTGTCCATCTCAGCCCCCAGAGT 3120  
 QY 3121 ATATCTGTCTTGGGAATCTCACACAGAACTCAGGAGCACCCCTGCCTGAGCTAAGG 3180  
 Db 3121 ATATCTGTCTTGGGAATCTCACACAGAACTCAGGAGCACCCCTGCCTGAGCTAAGG 3180  
 QY 3181 GAGGTCTTATCTCTCA-GGGGGGGTTTAAAGTCCCGTTTGAATAATGTCTTATTTAT 3239  
 Db 3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAAGTCCCGTTTGAATAATGTCTTATTTAT 3240  
 QY 3240 TTAGCGGGTGAATATTTTATCTGTAAGTGAAGCAATCAGATATAATGTTTATGGTGAC 3299  
 Db 3241 TTAGCGGGTGAATATTTTATCTGTAAGTGAAGCAATCAGATATAATGTTTATGGTGAC 3300  
 QY 3300 AAAATTAAAGGCTTTCTTAT 3319  
 Db 3301 AAAATTAAAGGCTTTCTTAT 3320

RESULT 13

AAS64040

ID AAS64040 standard; cDNA; 4034 BP.

XX AC

XX AC

XX AC

DT 29-JAN-2002 (first entry)

XX Human prosate cDNA P553S splice variant #3.

DE

XX

Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.

XX Homo sapiens.

PN WO200173032-A2.

XX PD 04-OCT-2001.

XX PF 27-MAR-2001; 2001WO-US09919.

XX PR 27-MAR-2000; 2000US-0536857.

XX PR 09-MAY-2000; 2000US-0568100.

XX PR 12-MAY-2000; 2000US-0570737.

XX PR 13-JUN-2000; 2000US-0593793.

XX PR 27-JUN-2000; 2000US-0605783.

XX PR 10-AUG-2000; 2000US-0636215.

XX PR 29-AUG-2000; 2000US-0651236.

XX PR 06-SEP-2000; 2000US-0657279.

XX PR 02-OCT-2000; 2000US-0679426.

XX PR 10-OCT-2000; 2000US-0685166.

XX PA (CORI-) CORIXA CORP.

XX XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;

PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;

PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPT; 2001-639232/73.

XX New human prostate-specific polypeptides and polynucleotides useful for

PT the diagnosis and treatment of cancer, especially prostate cancer -

PT Claim 1; Page 483-484; 579pp; English.

XX The invention relates to isolated prostate-specific

CC polynucleotides, polypeptides, fusion proteins of the polypeptides,

CC antibodies raised against the polypeptides (or antigenic epitopes

CC derived from them) and antigen-presenting cells expressing the

CC polypeptides. The antibodies are useful for detecting the presence of

CC cancer, especially prostate cancer. The polypeptides, polynucleotides and

CC the antigen-presenting cells are useful for stimulating and/or expanding

CC T cells specific for a tumour protein, and for inhibiting the development

CC of cancer especially prostate cancer. Compositions comprising the

CC polynucleotide and/or polypeptide are useful for stimulating an immune

CC response, and for treating cancer. The oligonucleotide is useful for

CC detecting cancer. The present sequence is a prostate specific

CC polynucleotide of the invention.

XX SQ Sequence 4034 BP; 721 A; 1226 C; 1141 G; 946 T; 0 other;

Query Match 75.8%; Score 2585.4; DB 22; Length 4034;

Best Local Similarity 82.4%; Pred. No. 0;

Matches 3325; Conservative 0; Mismatches 1; Indels 709; Gaps 3;

QY 4 AACGAGCTGCACGGCTGGCTCCGGGTGACAGCCGCGCTCGGCCAGATCTGAGTG 63

Db 1 AACGAGCTGCACGGCTGGCTCCGGGTGACAGCCGCGCTCGGCCAGATCTGAGTG 60

QY 64 ATGAGACGTGTCCTCCACTGAGGTGCCACAGCAGCAGGTGTTGAGCATGGCTGAGAAG 123

Db 61 ATGAGACGTGTCCTCCACTGAGGTGCCACAGCAGCAGGTGTTGAGCATGGCTGAGAAG 120

QY 124 CTGGACCGCCACCAAGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGGCAGTTGGC 183

Db 121 CTGGACCGCCACCAAGGCTGGCAGAAATGGGCGCTGGCTGATTCCTAGGCAGTTGGC 180

QY 184 GGCAGCAAGGAGGAGAGCCCGCAGCTTCTGGAGCAGAGCCGACGAGTCTTGAG 243

Db 181 GGCAGCAAGGAGGAGAGCCCGCAGCTTCTGGAGCAGAGCCGACGAGTCTTGAG 240

QY 244 TGCCTGAACGGCCCGCTGAGCCCTACCGCCCTGGCCCACTATGTTCCAGAGGCTGTGGGT 303

Db 241 TGCCTGAACGGCCCGCTGAGCCCTACCGCCCTGGCCCACTATGTTCCAGAGGCTGTGGGT 300



Qy	304	GAGCCGCTGCTGGGGCAACCGGAAGCCCAAGCTCTTCTGGTCAACCTGCTAACCTTTGG	363
Db	301		
		GAGCCGCTGCTGGGGCAACCGGAAGCCCAAGCTCTTCTGGTCAACCTGCTAACCTTTGG	360
Qy	364	CCTGGAGGTGTTTGGCCGCAAGCATCACTATATGTCGCGCTCTGCTGCTGGAAGTGG	423
Db	361		
		CCTGGAGGTGTTTGGCCGCAAGCATCACTATATGTCGCGCTCTGCTGCTGGAAGTGG	420
Qy	424	GGTAGAGGAAGTTTCATGACCATGTGTGGCATTTGGTCCAGTGTGGGCTTGCTGCTG	483
Db	421		
		GGTAGAGGAAGTTTCATGACCATGTGTGGCATTTGGTCCAGTGTGGGCTTGCTGCTG	480
Qy	484	TGTCGCCCTCCTAGGCTCAGCCAGTACCACTGGCGTGAGCGTATGGCCGCGCCGGCC	543
Db	481		
		TGTCGCCCTCCTAGGCTCAGCCAGTACCACTGGCGTGAGCGTATGGCCGCGCCGGCC	540
Qy	544	CTTCATCTGGGCACCTGTCTTTGGGCATCCCTGCTGAGCCTCTTTCTCATCCCAAGGCCG	603
Db	541		
		CTTCATCTGGGCACCTGTCTTTGGGCATCCCTGCTGAGCCTCTTTCTCATCCCAAGGCCG	600
Qy	604	CTGGCTAGCAGGGCTGCTGTGCCCGATCCAGGCCCTTGAGCTGGCACTGCTCATCCT	663
Db	601		
		CTGGCTAGCAGGGCTGCTGTGCCCGATCCAGGCCCTTGAGCTGGCACTGCTCATCCT	660
Qy	664	GGCGTGGGGCTGCTGGACTTCTGTGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT	723
Db	661		
		GGCGTGGGGCTGCTGGACTTCTGTGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT	720
Qy	724	CTCTGACCTCTTCGGGACCCGGACCACTGTGCGCAGGCCCTACTCTGTCTATGCCCTTCAT	783
Db	721		
		CTCTGACCTCTTCGGGACCCGGACCACTGTGCGCAGGCCCTACTCTGTCTATGCCCTTCAT	780
Qy	784	GATCAGTCTTGGGGCTGCTGGCTACCTTCCTGCCCTGCCATTTGACTGGGACACCACTGC	843
Db	781		
		GATCAGTCTTGGGGCTGCTGGCTACCTTCCTGCCCTGCCATTTGACTGGGACACCACTGC	840
Qy	844	CCTGGCCCCCTACTGGGCACCCAGGAGAGTGCCTCTTTGGCTGCTCACCTCATCTTT	903
Db	841		
		CCTGGCCCCCTACTGGGCACCCAGGAGAGTGCCTCTTTGGCTGCTCACCTCATCTTT	900
Qy	904	CCTCAGCTGGTAGCAGCACACTGCTGTGTGGCTGAGGAGGACGCTGGGCCCCACCGA	963
Db	901		
		CCTCAGCTGGTAGCAGCACACTGCTGTGTGGCTGAGGAGGACGCTGGGCCCCACCGA	960
Qy	964	GCCAGCAGAAGGCTGTGCGCCCCCTCTTGTGCGCCCCACTGCTGTGCCATTCGCGGGCCG	1023
Db	961		
		GCCAGCAGAAGGCTGTGCGCCCCCTCTCTTGTGCGCCCCACTGCTGTGCCATTCGCGGGCCG	1020
Qy	1024	CTTTGGCTTTCCGAAACCTGGGGCCCTGCTTCCCGGGCTGCACAGTGTGCTGCCGAT	1083
Db	1021		
		CTTTGGCTTTCCGAAACCTGGGGCCCTGCTTCCCGGGCTGCACAGTGTGCTGCCGAT	1080
Qy	1084	GCCCGCACCTGCGCGGCTCTTCGTGGCTGAGCTGTGCACTGGATGCACTCATGAC	1143
Db	1081		
		GCCCGCACCTGCGCGGCTCTTCGTGGCTGAGCTGTGCACTGGATGCACTCATGAC	1140
Qy	1144	CTTCAGCTGTTTACACGATTTCGTGGCGAGGGCTGTACACAGGCGTGCCCAAGAGC	1203
Db	1141		
		CTTCAGCTGTTTACACGATTTCGTGGCGAGGGCTGTACACAGGCGTGCCCAAGAGC	1200
Qy	1204	TGAGCCGGCACCGAGCCCGGAGACACTATGATGA-----	1239
Db	1201		
		TGAGCCGGCACCGAGCCCGGAGACACTATGATGAAGGTGAAGCCCTGGCAGCCAGCAG	1260
Qy	1240	-----	1239
Db	1261	AGGCTGCTGTGGAGCGGCCACAGAGACGACACTCGGGGCTGTGCTGGGCTGGTGCC	1320
Qy	1240	-----	1239
Db	1321	TCGCCATCCCTGGCCCCGACTTCTCTGTCAAGAAAGTGGGATGGAGCCCATCTGCATACA	1380

Db	2461	 TCGCCATCCCTGGGATAGTGCCTCCTGCTGCCAGGTGCCCATCCCTGTTTATGGGCT	2520
Qy	1818	CCATTGTCCAGCTCAGCAGCTGTGCACGTGCCTATATGGTGTCTGCCGAGGCGCTGGGTC	1877
Db	2521	CCATTGTCCAGCTCAGCAGCTGTGCACGTGCCTATATGGTGTCTGCCGAGGCGCTGGGTC	2580
Qy	1878	TGTCGCCCATTTTACTTTTGCTTACACAGGTAGTATTTTGACAAGAGCGACTTTGGCCAAATACT	1937
Db	2581	TGTCGCCCATTTTACTTTTGCTTACACAGGTAGTATTTTGACAAGAGCGACTTTGGCCAAATACT	2640
Qy	1938	CAGCGTAGAAAACTTCCAGCACATTTGGGTGGAGGCCCTGCCCTACATGGGTGCCAGCTCC	1997
Db	2641	CAGCGTAGAAAACTTCCAGCACATTTGGGTGGAGGCCCTGCCCTACATGGGTGCCAGCTCC	2700
Qy	1998	CCGCTCCTGTTAGCCCATCGGGGCTCGCGGCTGGCCGACAGTTTCTGTTGCTGCCAAAG	2057
Db	2701	CCGCTCCTGTTAGCCCATCGGGGCTCGCGGCTGGCCGACAGTTTCTGTTGCTGCCAAAG	2760
Qy	2058	TAATGTGGCTCTCTGCTGCCACCCCTGTGCTGTGAGGTGCGTAGCTGCACAGCTGGGGC	2117
Db	2761	TAATGTGGCTCTCTGCTGCCACCCCTGTGCTGTGAGGTGCGTAGCTGCACAGCTGGGGC	2820
Qy	2118	TGGGGGCTCCCTCTCCTCTCTCCCAAGTCTCTAGGGCTGCCCTACCTGGAGGCTTCCAAG	2177
Db	2821	TGGGGGCTCCCTCTCCTCTCTCCCAAGTCTCTAGGGCTGCCCTACCTGGAGGCTTCCAAG	2880
Qy	2178	GGGTTTTCAGCTCGGACTTATACAGGAGGCCAAGAGGCTCCATCAGCTGGAATCGGG	2237
Db	2881	GGGTTTTCAGCTCGGACTTATACAGGAGGCCAAGAGGCTCCATCAGCTGGAATCGGG	2940
Qy	2238	GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGGCTCCTTASTGTGAGACACA	2297
Db	2941	GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGGCTCCTTASTGTGAGACACA	3000
Qy	2298	CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACHTACGTACCTGGTTTCCCATCTCTAAGC	2357
Db	3001	CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACHTACGTACCTGGTTTCCCATCTCTAAGC	3060
Qy	2358	CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGATGGGAGTTTCTAGGATGAACAAC	2417
Db	3061	CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGATGGGAGTTTCTAGGATGAACAAC	3120
Qy	2418	TCCTCCATGGGATTTGAACATATG--ACTTATTTAGGGGAAGAGTCTCTGAGGGGCAAC	2475
Db	3121	TCCTCCATGGGATTTGAACATATGAAGTTATTTGTAGGGGAAGAGTCTCTGAGGGGCAAC	3180
Qy	2476	ACACAAGAACCGTCCCTCAGCCACAGCACTGTCTTTTGGTGTATCCACCCCTCT	2535
Db	3181	ACACAAGAACCGTCCCTCAGCCACAGCACTGTCTTTTGGTGTATCCACCCCTCT	3240
Qy	2536	TACCTTTATCAGGATGTGGCCCTGTGGTCTTCTTGTGGCATCACAGACACAGGCAAT	2595
Db	3241	TACCTTTATCAGGATGT--GCCGTGGTCCCTCTCTGTGCCATCACAGACACAGGCAAT	3299
Qy	2596	TTAAATATTTAAGTATTTATTTAACAAAGTAGAAGGAATCCATTGCTAGCTTTTCTGT	2655
Db	3300	TTAAATATTTAAGTATTTATTTAACAAAGTAGAAGGAATCCATTGCTAGCTTTTCTGT	3359
Qy	2656	GTGGGTCTTAATATTTTGGGTAGGTGGGGATCCCAACAATCAGTCCCTTGAGATAG	2715
Db	3360	GTGGGTCTTAATATTTTGGGTAGGTGGGGATCCCAACAATCAGTCCCTTGAGATAG	3419
Qy	2716	CTGCTATTGGGCTGATCATTCACGAATCTTCTCTCGGGGCTGGCCCCCAAAAT	2775
Db	3420	CTGCTATTGGGCTGATCATTCACGAATCTTCTCTCGGGGCTGGCCCCCAAAAT	3479
Qy	2776	GCCTAACCCAGGACCTTGGAAATTTCTACTCATCCCAAAATGATAATTCAAAATGCTGTAC	2835
Db	3480	GCCTAACCCAGGACCTTGGAAATTTCTACTCATCCCAAAATGATAATTCAAAATGCTGTAC	3539
Qy	2836	CCAAGGTTAGGGTGTGAAGGAAGGTAGAGGGTGGGGCTTCAGGCTTCAACGGCTTCCCT	2895

Db	3540	CCAAGTTAGGGTGTGAAGAAAGGTAGAGGGTGGGGCTTCAGGTCTCAACAGGCTTCCCT	3599
Qy	2896	AACACACCCCTCTCTCTTTGGCCAGCCTGGTTCCCCCACTTCCACATCCCTCTACTCTC	2955
Db	3600	AACACACCCCTCTCTCTTTGGCCAGCCTGGTTCCCCCACTTCCACATCCCTCTACTCTC	3659
Qy	2956	TCTAGGACTGGGCTGATGAAGGCACCTGCCCAAAATTTCCCTACCCCACTTTCCCTA	3015
Db	3660	TCTAGGACTGGGCTGATGAAGGCACCTGCCCAAAATTTCCCTACCCCACTTTCCCTA	3719
Qy	3016	CCCCCACTTTCCCAACAGGCTCCACAACCTGTTTGGAGCTACTGAGCAGCACAAGCA	3075
Db	3720	CCCCCACTTTCCCAACAGGCTCCACAACCTGTTTGGAGCTACTGAGCAGCACAAGCA	3779
Qy	3076	CAAAAGTCGGTTTCCCAAGCCTTTGTCATCTCAGCCGCCAGAGTATATCTGTCTTGGG	3135
Db	3780	CARAGTCGGTTTCCCAAGCCTTTGTCATCTCAGCCGCCAGAGTATATCTGTCTTGGG	3839
Qy	3136	GAATCTCACACAGAAATCAGGAGCAGCCCTGCCTGAGCTAAGGAGGCTTTATCTCTC	3195
Db	3840	GAATCTCACACAGAAATCAGGAGCAGCCCTGCCTGAGCTAAGGAGGCTTTATCTCTC	3899
Qy	3196	AGGGGGGTTTAACTGCGGTTTGCAATAATGTCGCTTATTTATAGCGGGTGAATAT	3255
Db	3900	AGGGGGGTTTAACTGCGGTTTGCAATAATGTCGCTTATTTATAGCGGGTGAATAT	3959
Qy	3256	TTTATACTGAAGTGAGCAATCAGAGTATATGTTTATGGTGACAAAATTAAGGCTTTC	3315
Db	3960	TTTATACTGAAGTGAGCAATCAGAGTATATGTTTATGGTGACAAAATTAAGGCTTTC	4019
Qy	3316	TTATATGTTTAAAAA 3330	
Db	4020	TTATATGTTTAAAAA 4034	
RESULT 14			
AAH93868			
ID	AAH93868	standard; cDNA: 4034 BP.	
XX	AAH93868;		
AC	AAH93868;		
XX	04-OCT-2001 (first entry)		
XX	P553S cDNA splice variant P553S-10.		
DE	Human: prostate cancer; prostate-specific; diagnosis; vaccine;		
KW	cystostatic; gene therapy; metastasis; ss.		
KW	Homo sapiens.		
XX	WO200151633-A2.		
XX	19-JUL-2001.		
XX	16-JAN-2001; 2001WO-US01574.		
PF	14-JAN-2000; 2000US-0483672.		
XX	(CORI-) CORIYA CORP.		
XX	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;		
PI	Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;		
PI	Wang A, Meagher MJ;		
XX	WPI: 2001-425873/45.		
DR	New polynucleotide encoding a prostate-specific protein, for		
XX	diagnosing, monitoring and treating prostate cancer in a patient and		
PT	for use in vaccines		
PT	Claim 1; Page 460-461; 543pp; English.		
XX	The present invention describes polynucleotide sequences (1) which encode		
XX			

prostate-specific proteins (II). (I) and (II) have cytostatic activity, and can be used in vaccine production and gene therapy. (I), (II), antibodies to (II), fusion proteins comprising (II), and isolated T cells prepared using (I) or (II) are used to treat cancer in a patient. (I) and the antibodies are also used in the detection of cancer in a patient. The cancer that is diagnosed or treated is particularly prostate cancer. (I) and (II) can be used in vaccines. The antibodies or (I) can be used for monitoring the progression of cancer in a patient. (I) and (II) can also be used to improve diagnostic and therapeutic methods for prostate cancer. They can indicate the level of metastasis as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to AAH01318 represent polynucleotide and amino acid sequences used in the exemplification of the present invention.

Query Match 75.88; Score 2585.4; DB 22; Length 4034;  
Best Local Similarity 82.48; Pred. No. 0;  
Matches 3325; Conservative 0; Mismatches 1; Indels 709; Gaps 3;

QY 4 AACGAGCTGTCACGCGTGGCTCCGGGTGACAGCCGCGCCCTCGGCCAGGATCTGAGTG 63  
DB 1 AACGAGCTGTCACGCGTGGCTCCGGGTGACAGCCGCGCCCTCGGCCAGGATCTGAGTG 60  
QY 64 ATGAGAGCTGTCGCCACTGAGTGCCGCCACAGCAGCAGAGTGTTGACATGGGCTGAGAAG 123  
DB 61 ATGAGAGCTGTCGCCACTGAGTGCCGCCACAGCAGCAGAGTGTTGACATGGGCTGAGAAG 120  
QY 124 CTGACCGGACCAAGAGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGGAGTTGGC 183  
DB 121 CTGACCGGACCAAGAGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGGAGTTGGC 180  
QY 184 GGCAGCAGGAGGAGGCGCCAGCTTCTGGAGCAGAGCCGAGCAGAAAGCTTCTGGAG 243  
DB 181 GGCAGCAGGAGGAGGCGCCAGCTTCTGGAGCAGAGCCGAGCAGAAAGCTTCTGGAG 240  
QY 244 TGCTGACGCGCCCTGAGCCCTACCCGCTGCCCCATATGGTCCAGAGGCTGTGGGT 303  
DB 241 TGCTGACGCGCCCTGAGCCCTACCCGCTGCCCCATATGGTCCAGAGGCTGTGGGT 300  
QY 304 GAGCGGCTGTGCGGCGACCGAAAGCCAGCTTCTGGTCAAGCTGTCAACCTTTGG 363  
DB 301 GAGCGGCTGTGCGGCGACCGAAAGCCAGCTTCTGGTCAAGCTGTCAACCTTTGG 360  
QY 364 CTTGAGAGTGTGTTGGCCGAGGATCACCCTATGTGCGCCCTCTGCTGTGGAAGTGG 423  
DB 361 CTTGAGAGTGTGTTGGCCGAGGATCACCCTATGTGCGCCCTCTGCTGTGGAAGTGG 420  
QY 424 GGTAGAGGAGAGTTTCAATGACCATGTTGCTGGGCTTGGTCCAGTGTGGCTGTGCTG 483  
DB 421 GGTAGAGGAGAGTTTCAATGACCATGTTGCTGGGCTTGGTCCAGTGTGGCTGTGCTG 480  
QY 484 TGTCCCGCTCCTAGGCTCAGCGATGACCACTGGCGTGGAGCTATGGCGCGCGCGGC 543  
DB 481 TGTCCCGCTCCTAGGCTCAGCGATGACCACTGGCGTGGAGCTATGGCGCGCGCGGC 540  
QY 544 CTTGATCTGGGCTGCTTGGGCTGCTGCTGAGCCTTTCTCATCCCAAGGCGCG 603  
DB 541 CTTGATCTGGGCTGCTTGGGCTGCTGCTGAGCCTTTCTCATCCCAAGGCGCG 600  
QY 604 CTGCTAGCAGGCTGTGTCGGGATCCAGGCCCTGGAGCTGGCATGCTCATCCT 663  
DB 601 CTGCTAGCAGGCTGTGTCGGGATCCAGGCCCTGGAGCTGGCATGCTCATCCT 660  
QY 664 GGGCGTGGGCTGTGAGCTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGT 723  
DB 661 GGGCGTGGGCTGTGAGCTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGT 720  
QY 724 CTCTGACCTCTTCCGGGACCGGACCACTGTCCGAGGCTACTGTCTATGCGCTTCA 783  
DB 721 CTCTGACCTCTTCCGGGACCGGACCACTGTCCGAGGCTACTGTCTATGCGCTTCA 780  
QY 784 GATCAGTCTTGGGGCTGCTGGGCTACCTCCTGCTGCCATGCTGGGACACCACTGC 843

DB 781 GATCAGTCTTGGGGCTGCTGGGTACCTCTCTGCTGCCATTGACTGGACACAGTGC 840  
QY 844 CTTGGCCCTTACCTTGGGACACAGGAGTGCCTCTTTGGCCCTGCTCACCTCATCTT 903  
DB 841 CTTGGCCCTTACCTTGGGACACAGGAGTGCCTCTTTGGCCCTGCTCACCTCATCTT 900  
QY 904 CTTACCTTCTGCTAGCAGCCACTCTGCTGGTGGCTGAGGAGGAGGCTGGGCCCCACCGA 963  
DB 901 CTTACCTTCTGCTAGCAGCCACTCTGCTGGTGGCTGAGGAGGAGGCTGGGCCCCACCGA 960  
QY 964 GGCAGAGAGGCTGTGGGCCCTCTCTTGTGCGCCCACTGCTCTCATGCGGCGCG 1023  
DB 961 GGCAGAGAGGCTGTGGGCCCTCTCTTGTGCGCCCACTGCTCTCATGCGGCGCG 1020  
QY 1024 CTTGGCTTTCGGGAACCTGGGCGCCCTGCTTCCCGGCTGACACAGCTGTGCTGCCGAT 1083  
DB 1021 CTTGGCTTTCGGGAACCTGGGCGCCCTGCTTCCCGGCTGACACAGCTGTGCTGCCGAT 1080  
QY 1084 GCCCGCACCTTSCGCCGCGCTCTTCTGGCTGAGCTGTCCAGCTGATGGACTCATGAC 1143  
DB 1081 GCCCGCACCTTSCGCCGCGCTCTTCTGGCTGAGCTGTCCAGCTGATGGACTCATGAC 1140  
QY 1144 CTTACAGCTGTTTACAGGATTTCTGGGCGAGGCTGTACAGGCGGTGCCAGAGC 1203  
DB 1141 CTTACAGCTGTTTACAGGATTTCTGGGCGAGGCTGTACAGGCGGTGCCAGAGC 1200  
QY 1204 TCAGCGGCGCACCGAGGCGCGGAGACTATGATGA 1239  
DB 1201 TCAGCGGCGCACCGAGGCGCGGAGACTATGATGAAGGTAAAGGCTTGGAGCCAGCAG 1260  
QY 1240 1239  
DB 1261 AGGCTGGTGTGGGAGCCGCCACAGACAGACTCTGGGGCTGTGCTGGGCTGTGCC 1320  
QY 1240 1239  
DB 1321 TCTCCATCTGCCCCGACTTCTGTGTCAGGAAAGTGGGATGGACCCCATCTGCAATACA 1380  
QY 1240 1239  
DB 1381 CGGCTTCTCATGGGTGTGAACATCTCTGCTTGGGTTTCAGGAAGGCTCTGGCTGCTC 1440  
QY 1240 1239  
DB 1441 TAGGAGTCTGATCAGAGTCTGTGCCCCAGTTTGACAGAAAGGAGGAGCTTATCAA 1500  
QY 1240 1239  
DB 1501 AGTCTAGAGGAGTGGAGAGTTAAGGCTGGAATTCAGATCTGCTGTTCAGCCGCGAG 1560  
QY 1240 1239  
DB 1561 TGTGCCCTCTGCTCCCCCAAGACTTTCCTCAATATCTCACAGGCGCTTCCAGCTCAG 1620  
QY 1240 1239  
DB 1621 CGTCTAGAGGCTTCTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCT 1680  
QY 1240 1239  
DB 1681 GTCCTCACAGTGAAGTCCCGAGGAACCTTCAGACTACCTTCTCTCTGCTTTCAGCAAG 1740  
QY 1240 1239  
DB 1741 GSGCTTGCCACATCTCTGAGGGTCACTGGAAGAACTAGACTCCCATTCCTAGAGTA 1800  
QY 1240 1239  
DB 1801 GAAAGGGAAGGCTGTGGGAGCAGGCTGTGTCACAGCAGGCTCTCTGTCAGCAGGTAC 1860  
QY 1240 1239



## RESULT 15

ABL95411

ID ABL95411 standard; cDNA; 4034 BP.

XX AC ABL95411;

XX 19-JUL-2002 (first entry)

XX DE Human P553S splice variant SEQ ID NO 704.

XX XX Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;

XX KW gene therapy; gene; ss.

XX OS Homo sapiens.

XX XX US2002022248-A1.

PN PD 21-FEB-2002.

XX PF 12-JAN-2001; 2001US-0759143.

XX PR 25-FEB-1997; 97US-0806099.

XX PR 01-AUG-1997; 97US-0904804.

XX PR 09-FEB-1998; 98US-0020956.

XX PR 25-FEB-1998; 98US-0030607.

XX PR 14-JUL-1998; 98US-0115453.

XX PR 23-SEP-1998; 98US-0159812.

XX PR 15-JAN-1999; 98US-0232149.

XX PR 09-APR-1999; 99US-0288946.

XX PR 13-JUL-1999; 99US-0352616.

XX PR 12-NOV-1999; 99US-0439313.

XX PR 18-NOV-1999; 99US-0443686.

XX PR 14-JAN-2000; 2000US-0483672.

XX PR 27-MAR-2000; 2000US-0536857.

XX PR 09-MAY-2000; 2000US-0568100.

XX PR 12-MAY-2000; 2000US-0570737.

XX PR 13-JUN-2000; 2000US-0593793.

XX PR 27-JUN-2000; 2000US-0605783.

XX PR 10-AUG-2000; 2000US-0636215.

XX PR 29-AUG-2000; 2000US-0651236.

XX PR 06-SEP-2000; 2000US-0657279.

XX PR 02-OCT-2000; 2000US-0679426.

XX PR 10-OCT-2000; 2000US-0685166.

XX (XUJJ/) XU J.

PA (DILL/) DILLON D C.

PA (MITC/) MITCHAM J L.

PA (HARL/) HARLOCKER S L.

PA (JIAN/) JIANG Y.

PA (KALO/) KALOS M D.

PA (FANG/) FANGER G R.

PA (RETT/) RETTER M W.

PA (STOL/) STOLK J A.

PA (DAYC/) DAY C H.

PA (VEDV/) VEDVICK T S.

PA (CART/) CARTER D.

PA (LISX/) LI S X.

PA (WANG/) WANG A.

PA (SKEI/) SKEIKY Y A W.

PA (HEPL/) HEPLER W T.

PA (HEND/) HENDERSON R A.

XX Xu J, Dillon DC, Harlocker SL, Jiang Y, Kalos MD;

PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;

PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2002-255649/30.

XX New prostate-specific polynucleotides for diagnosing and treating

PT diseases, in particular prostate cancer, and as markers for the

PT progression of cancer

XX

PS Claim 1; SEQ ID NO 704; 87pp; English.

XX The present invention provides prostate-specific coding sequences and  
 CC their encoded proteins. These can be used in the diagnosis and treatment  
 CC of cancers, particularly prostate cancer. The present sequence is a cDNA  
 CC described in the invention.

XX Sequence 4034 BP; 721 A; 1226 C; 1141 G; 946 T; 0 other;

Query Match 75.8%; Score 2585.4; DB 24; Length 4034;

Best Local Similarity 82.4%; Pred. No. 0;

Matches 3325; Conservative 0; Mismatches 1; Indels 709; Gaps 3;

QY 4 AACCAGCCTGCACGGCGTGGCTCCGGGTGACAGCGCGCGCTCGGCCAGGATCTGAGTG 63

DB 1 AACCAGCCTGCACGGCGTGGCTCCGGGTGACAGCGCGCGCTCGGCCAGGATCTGAGTG 60

QY 64 ATGAGACGTGTCCCCACTGAGTGGCCCCACACAGCAGAGGTGTGAGCATGGGCTGAGAG 123

DB 61 ATGAGACGTGTCCCCACTGAGTGGCCCCACACAGCAGAGGTGTGAGCATGGGCTGAGAG 120

QY 124 CTGGACCGGCACCAAGGGCTGGCAGAAATGGCGCTGGCTGATTCTCTAGGACGTTGGC 183

DB 121 CTGGACCGGCACCAAGGGCTGGCAGAAATGGCGCTGGCTGATTCTCTAGGACGTTGGC 180

QY 184 GGCAGCAAGGAGGAGAGCGCGCAGCTTCTGGAGCAGAGCCGAGAGCAAGCACTTCTGGAG 243

DB 181 GGCAGCAAGGAGGAGAGCGCGCAGCTTCTGGAGCAGAGCCGAGAGCAAGCACTTCTGGAG 240

QY 244 TGCCTGAACGGCCCCCTGAGCCCTAGCCGCTGGCCCACTATGTTCCAGAGGCTGTGGGT 303

DB 241 TGCCTGAACGGCCCCCTGAGCCCTAGCCGCTGGCCCACTATGTTCCAGAGGCTGTGGGT 300

QY 304 GAGCCGCTCTGCGGCACCCGAAAGCCAGCTTCTGTGGTCAACCTGCTAACCTTTGG 363

DB 301 GAGCCGCTCTGCGGCACCCGAAAGCCAGCTTCTGTGGTCAACCTGCTAACCTTTGG 360

QY 364 CTTGAGAGTGTGTTTGGCGCAGGATACCTATGTGCCGCTCTGCTGCTGGAAGTGGG 423

DB 361 CTTGAGAGTGTGTTTGGCGCAGGATACCTATGTGCCGCTCTGCTGCTGGAAGTGGG 420

QY 424 GGTAGAGGAGAGTTCATGACCATGTGCTGGCATTGCTCCAGTGTGGGCTGTGCTG 483

DB 421 GGTAGAGGAGAGTTCATGACCATGTGCTGGCATTGCTCCAGTGTGGGCTGTGCTG 480

QY 484 TGTCCGCTCTAGGCTCAGCCAGTACCACCTGGCGTGGACGCTATGCGCCGCGCGGCC 543

DB 481 TGTCCGCTCTAGGCTCAGCCAGTACCACCTGGCGTGGACGCTATGCGCCGCGCGGCC 540

QY 544 CTTTCATCTGGGCACTGCTTGGGCATCTGCTGAGCCTCTTTCTCATCCCAAGGCGCGG 603

DB 541 CTTTCATCTGGGCACTGCTTGGGCATCTGCTGAGCCTCTTTCTCATCCCAAGGCGCGG 600

QY 604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCAGGCCCTGGAGCTGGCACTGCTCATCT 663

DB 601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCAGGCCCTGGAGCTGGCACTGCTCATCT 660

QY 664 GGGCGTGGGCTGCTGAGCTTCTGTGGCAGGTGCTTCACTCCACTGGAGGCCCTGCT 723

DB 661 GGGCGTGGGCTGCTGAGCTTCTGTGGCAGGTGCTTCACTCCACTGGAGGCCCTGCT 720

QY 724 CTCTGACCTCTTCCGGGACCCGACACCTGTCGCCAGGCTACTCTGTCTATGCTCTCAT 783

DB 721 CTCTGACCTCTTCCGGGACCCGACACCTGTCGCCAGGCTACTCTGTCTATGCTCTCAT 780

QY 784 GATCAGTCTTGGGGCTGCTGGGCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843

DB 781 GATCAGTCTTGGGGCTGCTGGGCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

QY 844 CTTGGCCCCCTACTTGGCCACCCAGGAGGAGTGGCTTTTGGGCTGCTCACCTCATCTT 903

DB 841 CTTGGCCCCCTACTTGGCCACCCAGGAGGAGTGGCTTTTGGGCTGCTCACCTCATCTT 900

QY	904	CCTCACCTCGGTAGACGCACACTGCTGGTGGCTGAGAGCAGCGCTGGGCCCCACCGA	963
Db	901	CCTCACCTCGGTAGACGCACACTGCTGGTGGCTGAGAGCAGCGCTGGGCCCCACCGA	960
QY	964	GCCAGCAGAAGGCGTGTCGGCCCCCTCTGTGCGCCCCACATGCTGTCCATGCCGGGCCG	1023
Db	961	GCACAGAAAGGCGTGTCGGCCCCCTCTGTGCGCCCCACATGCTGTCCATGCCGGGCCG	1020
QY	1024	CTTTGGCTTTCCGSAACCTTGGCGCCCTGCTTCCC CGGCTGCACCAAGCTGTGTCGCCGAT	1083
Db	1021	CTTTGGCTTTCCGSAACCTTGGCGCCCTGCTTCCC CGGCTGCACCAAGCTGTGTCGCCGAT	1080
QY	1084	GCCCGCACCTTGGCGCGGCTCTTGTGGCTGAGCTGTGCGAGCTGTGATGCGACTCATGAC	1143
Db	1081	GCCCGCACCTTGGCGCGGCTCTTGTGGCTGAGCTGTGCGAGCTGTGATGCGACTCATGAC	1140
QY	1144	CTTCAGCTGTTTTACAGGATTTCTGGCGAGGGCTGTACCAAGGCTGTGTCGCCGAT	1203
Db	1141	CTTCAGCTGTTTTACAGGATTTCTGGCGAGGGCTGTACCAAGGCTGTGTCGCCGAT	1200
QY	1204	TGAGCGGCGACCGAGGCGCGGAGACACTATGATGA	1239
Db	1201	TGAGCGGCGACCGAGGCGCGGAGACACTATGATGAAGGTAAGGCCTTGGCGCCAGCAG	1260
QY	1240	-----	1239
Db	1261	AGGCTGGTGGAGCGCGCCACAGAGACACACTCGGGGCTGTGTGGGCTGTGTC	1320
QY	1240	-----	1239
Db	1321	TCTCCATCTGSCCGGACTTCTGTGTCAGAAAGTGGGGATGGACCCCATCTGATACA	1380
QY	1240	-----	1239
Db	1381	CGGCTTCTCATGGGTGGAACATCTGTCTTGGGTTTCAGGAAGGCTTGGCTGCTC	1440
QY	1240	-----	1239
Db	1441	TAGAGTCTGATCAGAGTGGTTGCCCGGCTTTGACAGAAAGGAGGAGCTATTCAA	1500
QY	1240	-----	1239
Db	1501	AGTCTAGAGGAGTGGAGAGTTAAGGCTGGATTCAGATCTGCTGGTCCAGCGCGAG	1560
QY	1240	-----	1239
Db	1561	TGTGCCCTCTGCTCCCCAACGACTTTCAAAATACTCACAGCGCTTCCAGCTCAGG	1620
QY	1240	-----	1239
Db	1621	CGTCTAGAAGCGCTTTGAAGCCTATGCCAGCTGTCTTTGTPTCCCTCTCACCGGCT	1680
QY	1240	-----	1239
Db	1681	GTCCCTCAGAGTGCAGACTCCAGGAAACCTTCAGACTACCTCTCTGCTGCTTCCAGCAGG	1740
QY	1240	-----	1239
Db	1741	GGCGTTGCCACATTTCTGAGGGTCACTGGAAGAACTTAGACTTCCCATTTGCTAGAGGTA	1800
QY	1240	-----	1239
Db	1801	GAAAGGGAAGGTGCTGGGAGCAGGCGTGGTCCACAGCAGGTCTCTGTCAGCAGGTAC	1860
QY	1240	-----	1239
Db	1861	CTGTGTTCCGCTTCTCATCTCCCTGAGACTGCTCGACCGCTTCCCTCCAGGCTCTGT	1920
QY	1240	-----AGGCGTTCGGATGGCAGCGCTGGGCGTGTCTCTGCGT	1277
Db	1921	CTGATGGCGGCTCTCCCTCTGCAGGCGTTCGGATGGCAGCGCTGGGCGTGTCTCTGAGT	1980
QY	1278	GGCCATCTCCCTGCTCTCTCTGCTGTCATGAGCGGCTGGTGGCAGGATTCGGCAGCTC	1337



Db 3061 CCCTTAACCTGCAGCTTCGTTTAATGATAGCTCTTGCATGGAGATTTCTAGGATGAACAC 3120  
QY 2418 TCCTCCATGGATTTGAACATATG--ACTTATTGTAGGGAAGAGTCTCTGAGGGCAAC 2475  
Db 3121 TCCTCCATGGATTTGAACATATGAAAGTTATTTGTAGGGAAGAGTCTCTGAGGGCAAC 3180  
QY 2476 ACACAAGAACAGGTTCCTCAGCCACAGACACTGTCTTTTGTGTATCCACCCCTCT 2535  
Db 3181 ACACAAGAACAGGTTCCTCAGCCACAGACACTGTCTTTTGTGTATCCACCCCTCT 3240  
QY 2536 TACCTTTTATCAGGATGGCTGTCTGCTCTCTGTTGCCATCACAGACACAGGCAT 2595  
Db 3241 TACCTTTTATCAGGATGT--GCTGTGTGCTCTCTGTTGCCATCACAGACACAGGCAT 3299  
QY 2596 TTAATATTTTAACTTATTTTATTTTAAACAAAGTAGAAGGAATCCATGCTAGCTTTTCTGT 2655  
Db 3300 TTAATATTTTAACTTATTTTATTTTAAACAAAGTAGAAGGAATCCATGCTAGCTTTTCTGT 3359  
QY 2656 GTTGTGTCTTAATATTTGGGTAGGTGGGGATCCCAACAAATCAGGTCCCTGAGATAG 2715  
Db 3360 GTTGTGTCTTAATATTTGGGTAGGTGGGGATCCCAACAAATCAGGTCCCTGAGATAG 3419  
QY 2716 CTGTCATTTGGCTGATCATTTCCAGAAATCTTCTCTCTGGGTCTGGCCCCCAAAAT 2775  
Db 3420 CTGTCATTTGGCTGATCATTTCCAGAAATCTTCTCTCTGGGTCTGGCCCCCAAAAT 3479  
QY 2776 GCTTAACCCAGGACCTTGGAAATTTACTATCCCAATGATATTCAAATGCTGTAC 2835  
Db 3480 GCTTAACCCAGGACCTTGGAAATTTACTATCCCAATGATATTCAAATGCTGTAC 3539  
QY 2836 CCAAGGTAGGTGTGAAGAAAGTACAGGTGGGCTTCAGTCTCAACGGCTTCCT 2895  
Db 3540 CCAAGGTAGGTGTGAAGAAAGTACAGGTGGGCTTCAGTCTCAACGGCTTCCT 3599  
QY 2896 AACCACCCCTCTCTCTTTGGCCAGCCCTGTTCCTCCCACTTCCACTCCCTCTACTCTC 2955  
Db 3600 AACCACCCCTCTCTCTTTGGCCAGCCCTGTTCCTCCCACTTCCACTCCCTCTACTCTC 3659  
QY 2956 TCTAGACTGGCTGATGAAGCACTGCCCAAAATTTCCCTACCCCACTTTCCCTTA 3015  
Db 3660 TCTAGACTGGCTGATGAAGCACTGCCCAAAATTTCCCTACCCCACTTTCCCTTA 3719  
QY 3016 CCCCACACTTTCCCAACAGCTTCCCAACCTGTTTGGAGCTACTGCAGGACCAAGCA 3075  
Db 3720 CCCCACACTTTCCCAACAGCTTCCCAACCTGTTTGGAGCTACTGCAGGACCAAGCA 3779  
QY 3076 CAAAGTGGCTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGATATCTGCTTGG 3135  
Db 3780 CAAAGTGGCTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGATATCTGCTTGG 3839  
QY 3136 GAATCTCACAGAACTCAGAGCACCCCTGCTGAGCTAAGGGAGTCTTATCTCTC 3195  
Db 3840 GAATCTCACAGAACTCAGAGCACCCCTGCTGAGCTAAGGGAGTCTTATCTCTC 3899  
QY 3196 AGGGGGGTTTAAAGTCCGCTTTGCAATAATGTCGCTTATTTTATTTAGCGGGTGAATAT 3255  
Db 3900 AGGGGGGTTTAAAGTCCGCTTTGCAATAATGTCGCTTATTTTATTTAGCGGGTGAATAT 3959  
QY 3256 TTTATATCTAAGTGAAGCAATCAGAGTATAATGTTTATGTCACAAAATTAAGGCTTTC 3315  
Db 3960 TTTATATCTAAGTGAAGCAATCAGAGTATAATGTTTATGTCACAAAATTAAGGCTTTC 4019  
QY 3316 TTTATATGTTTAAAAA 3330  
Db 4020 TTTATATGTTTAAAAA 4034

## RESULT 16

ABK92217

ID ABK92217 standard; DNA; 2582 BP.

XX

AC ABK92217;

XX

DT 15-AUG-2002 (first entry)  
XX Prostate cancer-associated DNA sequence #103.  
DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;  
KW gene therapy; gene; ds.  
XX Mammalia.  
OS  
PN WO200230268-A2.  
XX  
PD 18-APR-2002.  
XX  
PF 12-OCT-2001; 2001WO-US32045.  
XX  
PR 13-OCT-2000; 2000US-0687576.  
PR 08-DEC-2000; 2000US-0733288.  
PR 08-DEC-2000; 2000US-0733742.  
PR 24-JAN-2001; 2001US-263957P.  
PR 16-MAR-2001; 2001US-276791P.  
PR 06-APR-2001; 2001US-276888P.  
PR 24-APR-2001; 2001US-281922P.  
PR 30-APR-2001; 2001US-286214P.  
PR 04-MAY-2001; 2001US-0847046.  
PR 04-MAY-2001; 2001US-288589P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
PA  
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezl P;  
PI  
XX WPI: 2002-471335/50.  
DR P-PSDB; ABG61900.  
XX  
PT Detecting a prostate cancer-associated transcript in a cell in a  
PT patient, useful for diagnosing prostate cancer (PC) or screening  
PT modulators of PC, by determining if prostate cancer-associated genes  
PT are expressed in a prostate tissue  
XX  
PS Claim 22; Page 386; 436pp; English.  
CC  
CC The present invention relates to methods of detecting a prostate  
CC cancer-associated transcript in a cell from a patient. The method  
CC comprises contacting a biological sample from the patient with  
CC prostate cancer-associated polynucleotides (designated PC genes) that  
CC selectively hybridise to a sequence that is at least 80% identical  
CC to them. The prostate cancer-associated polynucleotide sequences  
CC are differentially expressed in prostate tumour tissue or in  
CC prostate cancer and are derived from the tissues of various  
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).  
CC The methods of the invention are useful for diagnosing and treating  
CC prostate cancer in mammals. The prostate cancer-associated genes are  
CC useful for diagnosing or treating prostate cancer, as well as for  
CC identifying modulators of prostate cancer or agents that inhibit  
CC prostate cancer. The nucleic acid sequences are particularly useful  
CC in gene therapy, as a vaccine or in antisense applications.  
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide  
CC sequences.  
XX  
SQ Sequence 2582 BP; 407 A; 822 C; 785 G; 566 T; 2 other;

Query Match 72.9%; Score 2486.4; DB 24; Length 2582;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2499; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 2 GGAACCAAGCTGCACGGCTGGCTCCGGGTGACAGCGCGCGCTCGGCCAGGATCTAG 61  
Db 28 GGAACCAAGCTGCACGGCTGGCTCCGGGTGACAGCGCGCGCTCGGCCAGGATCTAG 87  
QY 62 TGATGAGACGTGTCCCACTGAGGTGCCCCACAGACAGAGGTGTTGAGCATGGCTGAGA 121  
Db 88 TGATGAGACGTGTCCCACTGAGGTGCCCCACAGACAGAGGTGTTGAGCATGGCTGAGA 147  
QY 122 AGCTGGACCGCCACCAAGGGCTGGCAGAAATGGCGCCTGGCTGATTCTTAGCGCATTTG 181







QY 1449 TCTAGCCCTCAGATCTCGCCCTACACACTGGCTCCCTCTACACCGGAGAGCAGG 1508  
 Db 1021 TCTAGCCCTCAGATCTCGCCCTACACACTGGCTCCCTCTACACCGGAGAGCAGG 1080  
 QY 1509 TGTTCCTGCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGGACACAGCCTGATGA 1568  
 Db 1081 TGTTCCTGCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGGACACAGCCTGATGA 1140  
 QY 1569 CAGCTTCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTTAATGGACACAGTGGGTCTG 1628  
 Db 1141 CAGCTTCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTTAATGGACACAGTGGGTCTG 1200  
 QY 1629 GAGGCACTGGCTGCTCCACCTCCACCCGCGCTCTCGGGGCTCTCGCTGCTGATGCTCT 1688  
 Db 1201 GAGGCACTGGCTGCTCCACCTCCACCCGCGCTCTCGGGGCTCTCGCTGCTGATGCTCT 1260  
 QY 1689 CGGTACGTGTGGTGGGTGAGCCACCGAGGCCAGGGTGGTTCGGGGCGGGGGCATCT. 1748  
 Db 1261 CGGTACGTGTGGTGGGTGAGCCACCGAGGCCAGGGTGGTTCGGGGCGGGGGCATCT 1320  
 QY 1749 GCTCGACCTCGCCATCTGGATAGTGTCTCTGCTGCTCCAGGTGGCCCATCCCTGT 1808  
 Db 1321 GCTCGACCTCGCCATCTGGATAGTGTCTCTGCTGCTCCAGGTGGCCCATCCCTGT 1380  
 QY 1809 TTATGGGTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCCTATATGCTGTCTGCCGAG 1868  
 Db 1381 TTATGGGTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCCTATATGCTGTCTGCCGAG 1440  
 QY 1869 GCTCGGTCTGTGCCATCTTACTTTGTGTACACAGGTAGTATTTGACAAAGACACTTGG 1928  
 Db 1441 GCTCGGTCTGTGCCATCTTACTTTGTGTACACAGGTAGTATTTGACAAAGACACTTGG 1500  
 QY 1929 CCAATATCTCAGCGTAGAAAACCTCCACACATTTGGGTGGAGGCCCTGCTCACTGGGT 1988  
 Db 1501 CCAATATCTCAGCGTAGAAAACCTCCACACATTTGGGTGGAGGCCCTGCTCACTGGGT 1560  
 QY 1989 CCCAGTCCCGCTCCTGTAGCCCATGGGCTCGGGCTCGGGCTGGCCGCTGCTGTGTG 2048  
 Db 1561 CCCAGTCCCGCTCCTGTAGCCCATGGGCTCGGGCTCGGGCTGGCCGCTGCTGTGTG 1620  
 QY 2049 CTGCCAAGTATGTGGCTCTCTGTGCCACCTGTCTGTGCTGAGGTGGCTAGCTGCACA 2108  
 Db 1621 CTGCCAAGTATGTGGCTCTCTGTGCCACCTGTCTGTGCTGAGGTGGCTAGCTGCACA 1680  
 QY 2109 GTGGGGCTGGGGCTGCTCTCTCTCCAGCTCTCTAGGCTGCTGACTGGAGG 2168  
 Db 1681 GTGGGGCTGGGGCTGCTCTCTCTCCAGCTCTCTAGGCTGCTGACTGGAGG 1740  
 QY 2169 CCTTCCAGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAGGGCTCCATGCACTG 2228  
 Db 1741 CCTTCCAGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAGGGCTCCATGCACTG 1800  
 QY 2229 GAATGGGGACTCTGCAGGTGGATTACCGGCTCAGGTTAAGAGCTAGCTCCTAGT 2288  
 Db 1801 GAATGGGGACTCTGCAGGTGGATTACCGGCTCAGGTTAAGAGCTAGCTCCTAGT 1860  
 QY 2289 TGAGACACACCTAGAGAGGGTTTGGAGCTGAATAAATCACTACCTGTTTCCCA 2348  
 Db 1861 TGAGACACACCTAGAGAGGGTTTGGAGCTGAATAAATCACTACCTGTTTCCCA 1920  
 QY 2349 TCTTAAGCCCTTAACCTGACGTCTGTTTAATGAGCTCTTGAATGGAGTTTCTAGG 2408  
 Db 1921 TCTTAAGCCCTTAACCTGACGTCTGTTTAATGAGCTCTTGAATGGAGTTTCTAGG 1980  
 QY 2409 ATGAACACTCTCCATGGATTGAACATATG - ACTTATTTGGAGGGAAGCTCTG 2466  
 Db 1981 ATGAACACTCTCCATGGATTGAACATATGAAAGTTATTTGAGGGGAAGGTCCTG 2040  
 QY 2467 AGGGCAACACAAAGAACAGGTCCCTCAGCCACACAGCACTGTCTTTTCTGATCCA 2526  
 Db 2041 AGGGCAACACAAAGAACAGGTCCCTCAGCCACACAGCACTGTCTTTTCTGATCCA 2100  
 QY 2527 CCCCCCTCTTACCTTTTATACAGGATGTGGCTGTGGTCTCTCTGTTGCCATCACAGAGA 2586

Db 2101 CCCCCCTCTACCTTTTATCAGGATGTGCCCTGTGTGCTCTCTGTGTGCCATCACAGAGA 2160  
 QY 2587 CACAGGATTTAAATATTTAACTTATTTAAACAAGTAGNAGGGAATCCATTCCTAG 2646  
 Db 2161 CACAGGATTTAAATATTTAACTTATTTAAACAAGTAGNAGGGAATCCATTCCTAG 2220  
 QY 2647 CTTTCTGTCTGTCTAAATTTGGGTAGGTGGGGATCCCAACAATCAGGTCC 2706  
 Db 2221 CTTTCTGTCTGTCTAAATTTGGGTAGGTGGGGATCCCAACAATCAGGTCC 2280  
 QY 2707 CTGAGATAGCTGTCTAATTTGGGTGATCATTTGCCAGAAATCTTCTCTCTGGGTCTGGCC 2766  
 Db 2281 CTGAGATAGCTGTCTAATTTGGGTGATCATTTGCCAGAAATCTTCTCTCTGGGTCTGGCC 2340  
 QY 2767 CCCCAAAATGCCCTAACCCAGGACCTTGGAAATTTCTACTATCCCAATGATATTTCCAAA 2826  
 Db 2341 CCCCAAAATGCCCTAACCCAGGACCTTGGAAATTTCTACTATCCCAATGATATTTCCAAA 2400  
 QY 2827 TGTCTTTACCAAGGTAGGTGTGTAAGGAAGGTAGAGGTGGGGCTTTCAGGTCTCAAC 2886  
 Db 2401 TGTCTTTACCAAGGTAGGTGTGTAAGGAAGGTAGAGGTGGGGCTTTCAGGTCTCAAC 2460  
 QY 2887 GGCTTCCCTAACCAACCCCTCTCTCTTGGGCCAGCCTGGTTCCCCCACTTCCACTCC 2946  
 Db 2461 GGCTTCCCTAACCAACCCCTCTCTCTTGGGCCAGCCTGGTTCCCCCACTTCCACTCC 2520  
 QY 2947 TCTACTCTCTAGGACTGGCTGTGTAAGGCACTGCCCAAAATTTCCCTACCCCAAC 3006  
 Db 2521 TCTACTCTCTAGGACTGGCTGTGTAAGGCACTGCCCAAAATTTCCCTACCCCAAC 2580  
 QY 3007 TTTTCCCTACCCCAACTTTCCCAACAGCTCCCAACACCTGTTGGAGCTACTGAGGA 3066  
 Db 2581 TTTTCCCTACCCCAACTTTCCCAACAGCTCCCAACACCTGTTGGAGCTACTGAGGA 2640  
 QY 3067 CCAGAAGCAAAAGTGGGTTCCTCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATCT 3126  
 Db 2641 CCAGAAGCAAAAGTGGGTTCCTCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATCT 2700  
 QY 3127 GTGCTTGGGAATCTCACAGAACTCAGGAGCACCCCTGCTGAGCTAAGGGAGGTC 3186  
 Db 2701 GTGCTTGGGAATCTCACAGAACTCAGGAGCACCCCTGCTGAGCTAAGGGAGGTC 2760  
 QY 3187 TTATCTCTCAGGGGGGTTTAAAGTGGCTTGAATTAATGCTGCTTATTTAGCGG 3246  
 Db 2761 TTATCTCTCAGGGGGGTTTAAAGTGGCTTGAATTAATGCTGCTTATTTAGCGG 2820  
 QY 3247 GGTGAATATTTTATCTAAGTCAAGCAATCAGAGTATTAATGTTATGTCACAAATTA 3306  
 Db 2821 GGTGAATATTTTATCTAAGTCAAGCAATCAGAGTATTAATGTTATGTCACAAATTA 2880  
 QY 3307 AAGGCTTCTTATATGTTTAAAAA 3330  
 Db 2881 AAGGCTTCTTATATGTTTAAAAA 2904

RESULT 18

AAH93867

ID AAH93867 standard; cDNA: 2904 BP.

XX AC AAH93867;

XX DT 04-OCT-2001 (first entry)

XX DE P553S cDNA splice variant p553S-12.

XX KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;

XX OS cytosolic; gene therapy; metastasis; ss.

XX PN Homo sapiens.

XX WO200151633-A2.

PD 19-JUL-2001.  
 XX  
 PF 16-JAN-2001; 2001WO-US01574.  
 XX  
 PR 14-JAN-2000; 2000US-0483672.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;  
 PI Wang A, Meagher MJ;  
 XX  
 DR WPI; 2001-425873/45.  
 XX  
 XX New polynucleotide encoding a prostate-specific protein, for  
 PT diagnosing, monitoring and treating prostate cancer in a patient and  
 PT for use in vaccines.  
 XX  
 PS Claim 1; Page 459-460; 543pp; English.  
 XX  
 XX The present invention describes polynucleotide sequences (I) which encode  
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
 CC and can be used in vaccine production and gene therapy. (I), (II),  
 CC antibodies to (II), fusion proteins comprising (II), and isolated  
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.  
 CC (I) and the antibodies are also used in the detection of cancer in a  
 CC patient. The cancer that is diagnosed or treated is particularly  
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
 CC (I) can be used for monitoring the progression of cancer in a patient.  
 CC (I) and (II) can also be used to improve diagnostic and therapeutic  
 CC methods for prostate cancer. They can indicate the level of metastasis  
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to  
 CC AAH01318 represent polynucleotide and amino acid sequences used in the  
 CC exemplification of the present invention.  
 XX  
 SQ Sequence 2904 BP; 542 A; 875 C; 773 G; 714 T; 0 other;

Query Match 64.4%; Score 2196.4; DB 22; Length 2904;  
 Best Local Similarity 88.2%; Pred. No. 0;  
 Matches 2560; Conservative 0; Mismatches 1; Indels 343; Gaps 2;

QY 1241 ----- 1240  
 Db 481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGCTTTGACAGAAAGGCGGA 540  
 QY 1241 ----- 1240  
 Db 541 GCTTATTAAAGTCTAGAGGAGTGGAGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600  
 QY 1241 ----- 1240  
 Db 601 CCAGCCGAGTGTGCCCTCTGCTCCCAACAGACTTTCCAAATAATCTCACCAGCCCTT 660  
 QY 1241 ----- 1240  
 Db 661 CCAGCTCAGGCGTCTCTAGAACGCTTTGAAGCCCTATGCCCAGCTGTCTTTGTGTTCCCTC 720  
 QY 1241 ----- 1240  
 Db 721 TCACCCGCTGTCTCTACAGCTGAGACTCCCAAGAAACCTTCAGACTACCTTCCTCTGCC 780  
 QY 1241 ----- -GGCCTTCGGATGGCAGCCTGGGCTGT 1268  
 Db 781 TTCAGCAAGGGGCGTTCGCCACACATTCCTGAGGGCGTTCGGATGGGAGCCTGGGGCTGT 840  
 QY 1269 TCCTGCAGTCGCCCATCTCCCTGTCTCTCTCTGTCTCATGACCGGCTGGTGCACGAT 1328  
 Db 841 TCCTGCAGTCGCCCATCTCCCTGTCTCTCTCTGTCTCATGACCGGCTGGTGCACGAT 900  
 QY 1329 TCGGCACTCCAGCAGTCTATTTGGCCAGTGTGCAGCTTTCCCTGTGGCTGCCGTGCCA 1388  
 Db 901 TCGGCACTCCAGCAGTCTATTTGGCCAGTGTGCAGCTTTCCCTGTGGCTGCCGTGCCA 960  
 QY 1389 CATGCTCTGTCCCAAGTGTGGCGCTGGTGCACAGCTTCAGCGGCCCTCACCGGTTCCACT 1448  
 Db 961 CATGCTCTGTCCCAAGTGTGGCGCTGGTGCACAGCTTCAGCGGCCCTCACCGGTTCCACT 1020  
 QY 1449 TCTACGCTCTGAGATCTGCTCCCTACACACTGGCCTCCCTTACCAACCGGAGAACAG 1508  
 Db 1021 TCTACGCTCTGAGATCTGCTCCCTACACACTGGCCTCCCTTACCAACCGGAGAACAG 1080  
 QY 1509 TGTTCCTGCCCAATACCGAGGGGACACTGGAGTGTCTAGCAGTGGAGACAGCTGATGA 1568  
 Db 1081 TGTTCCTGCCCAATACCGAGGGGACACTGGAGTGTCTAGCAGTGGAGACAGCTGATGA 1140  
 QY 1569 CCAGCTTCTTGCAGCCCTTAAGCCTGGAGCTCCCTTCCCTTAATGGACACGCTGGGTGCTG 1628  
 Db 1141 CCAGCTTCTTGCAGCCCTTAAGCCTGGAGCTCCCTTCCCTTAATGGACACGCTGGGTGCTG 1200  
 QY 1629 GAGGCAGTGGCTGCTCCCACTCCACCGGGCTGTGGGGGCTCTGCTGTGATGTCT 1688  
 Db 1201 GAGGCAGTGGCTGCTCCCACTCCACCGGGCTGTGGGGGCTCTGCTGTGATGTCT 1260  
 QY 1689 CCCTAGCTGTGGTGGGTGAGCCCAACGAGCCAGGCTGGTTCCGGGCCCGGCACTCT 1748  
 Db 1261 CCCTAGCTGTGGTGGGTGAGCCCAACGAGCCAGGCTGGTTCCGGGCCCGGCACTCT 1320  
 QY 1749 GCCTGACCTCGCCATCTCTGGATAGTGTCTTCCCTGTCCAGGTGGCCCATCCCTGT 1808  
 Db 1321 GCCTGACCTCGCCATCTCTGGATAGTGTCTTCCCTGTCCAGGTGGCCCATCCCTGT 1380  
 QY 1809 TTATGGCTCCATTTGTCAGCTCAGCCAGTCTGTCTACCTGCTATATGGTGTCTGCCGAG 1868  
 Db 1381 TTATGGCTCCATTTGTCAGCTCAGCCAGTCTGTCTACCTGCTATATGGTGTCTGCCGAG 1440  
 QY 1869 GCTGGTGTGGTGGGTGAGCCCAACGAGCCAGGCTGGTTCCGGGCCCGGCACTCT 1928  
 Db 1441 GCTGGTGTGGTGGGTGAGCCCAACGAGCCAGGCTGGTTCCGGGCCCGGCACTCT 1500  
 QY 1929 CCAATATCTACGCTAGAAACCTTCACGACATTTGGGTGGAGGGCTGCTCTACTGGGT 1988  
 Db 1501 CCAATATCTACGCTAGAAACCTTCACGACATTTGGGTGGAGGGCTGCTCTACTGGGT 1560  
 QY 1989 CCAGCTCCCCGCTCTCTGTAGCCCATAGGGGCTGCCGGGCTGGCGGCAGTTTCTGTG 2048

Db	2641	CCAGAAGCACAAGTGC	GGTGTTC	CCCAAGCCTTTGTCATCTCAGCCCCCAGAGTATATCT	2700
QY	3127	GTGCTTTGGGGAATCTC	CACACAGAAACTC	AGGAGCACCCCTGCCTGAGCTAAGGAGGTC	3186
Db	2701	GTGCTTTGGGGAATCTC	CACACAGAAACTC	AGGAGCACCCCTGCCTGAGCTAAGGAGGTC	2760
QY	3187	TTATCTCTCAGGGGGGGT	TTAAGTGCCGTTT	TGCAATAATGTCGCTTATTTATTTAGCGG	3246
Db	2761	TTATCTCTCAGGGGGGGT	TTAAGTGCCGTTT	TGCAATAATGTCGCTTATTTATTTAGCGG	2820
QY	3247	GGTGAATATTTTATAC	TGTAAGTGAGCAATC	AGAGTATTAATGTTATGTTGACAAAATTA	3306
Db	2821	GGTGAATATTTTATAC	TGTAAGTGAGCAATC	AGAGTATTAATGTTATGTTGACAAAATTA	2880
QY	3307	AAGGCTTCTTATATG	TTTAAAAA	3330	
Db	2881	AAGGCTTCTTATATG	TTTAAAAA	2904	
RESULT 19					
ABL95410					
XX	ID	ABL95410 standard; cDNA; 2904 BP.			
XX	AC	ABL95410;			
XX	DT	19-JUL-2002 (first entry)			
XX	XX	Human P53S splice variant SEQ ID NO 703.			
DE	DE	Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;			
XX	KW	gene therapy; gene; ss.			
XX	XX	Homo sapiens.			
XX	OS	US2002022248-A1.			
PN	PN	21-FEB-2002.			
XX	PD	12-JAN-2001; 2001US-0759143.			
XX	PF	25-FEB-1997; 97US-0806099.			
XX	PR	01-AUG-1997; 97US-0904804.			
PR	PR	09-FEB-1998; 98US-0020956.			
PR	PR	25-FEB-1998; 98US-0030607.			
PR	PR	14-JUL-1998; 98US-0115453.			
PR	PR	23-SEP-1998; 98US-0159812.			
PR	PR	15-JAN-1999; 99US-0232149.			
PR	PR	09-APR-1999; 99US-0288946.			
PR	PR	13-JUL-1999; 99US-0352616.			
PR	PR	12-NOV-1999; 99US-0439313.			
PR	PR	18-NOV-1999; 99US-0443686.			
PR	PR	14-JAN-2000; 2000US-0483672.			
PR	PR	27-MAR-2000; 2000US-0536857.			
PR	PR	09-MAY-2000; 2000US-0568100.			
PR	PR	12-MAY-2000; 2000US-0570737.			
PR	PR	13-JUN-2000; 2000US-0593793.			
PR	PR	27-JUN-2000; 2000US-0605783.			
PR	PR	10-AUG-2000; 2000US-0636215.			
PR	PR	29-AUG-2000; 2000US-0651236.			
PR	PR	06-SEP-2000; 2000US-0657279.			
PR	PR	02-OCT-2000; 2000US-0679426.			
PR	PR	10-OCT-2000; 2000US-0685166.			
XX	XX	(XUJJ/) XU J.			
PA	PA	(DILL/) DILLON D C.			
PA	PA	(MITC/) MITCHAM J L.			
PA	PA	(HARL/) HARLOCKER S L.			
PA	PA	(JIAN/) JIANG Y.			
PA	PA	(KALO/) KALOS M D.			
PA	PA	(FANG/) FANGER G R.			
PA	PA	(RETT/) RETTER M W.			
PA	PA	(STOL/) STOLK J A.			
PA	PA	(DAYC/) DAY C H.			

PA (VEDV/) VEDVICK T S.  
 PA (CART/) CARTER D.  
 PA (LISX/) LI S X.  
 PA (WANG/) WANG A.  
 PA (SKEI/) SKEIKY Y A W.  
 PA (HEPL/) HEPLER W T.  
 PA (HEND/) HENDERSON R A.  
 XX  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;  
 DR WPI; 2002-255649/30.  
 XX  
 XX New prostate-specific polynucleotides for diagnosing and treating  
 PT diseases, in particular prostate cancer, and as markers for the  
 PT progression of cancer.  
 XX  
 XX Claim 1; SEQ ID NO 703; 87pp; English.  
 PS  
 CC The present invention provides prostate-specific coding sequences and  
 CC their encoded proteins. These can be used in the diagnosis and treatment  
 CC of cancers, particularly prostate cancer. The present sequence is a CDNA  
 CC described in the invention.  
 XX  
 XX Sequence 2904 BP; 542 A; 875 C; 773 G; 714 T; 0 other;  
 SQ  
 Query Match 64.4%; Score 2196.4; DB 24; Length 2904;  
 Best Local Similarity 88.2%; Pred. No. 0;  
 Matches 2560; Conservative 0; Mismatches 1; Indels 343; Gaps 2;  
 QY 770 GTCATGCTTCATGATCAGTCTTGGGGGCTGCTGGGCTACCTCTCGCTGCGCATGAC 829  
 DB 1 GTCATGCTTCATGATCAGTCTTGGGGGCTGCTGGGCTACCTCTCGCTGCGCATGAC 60  
 QY 830 TGGGACACCACTGCTGGCCCTACCTCGGACCCAGGAGGAGTGCCTTTGGCCCTG 889  
 DB 61 TGGGACACCACTGCTGGCCCTACCTCGGACCCAGGAGGAGTGCCTTTGGCCCTG 120  
 QY 890 CTCACCTCATCTTCCTCACCCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAGGAGCG 949  
 DB 121 CTCACCTCATCTTCCTCACCCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAGGAGCG 180  
 QY 950 CTGGGCCCCCAGGACGAGAGGCTGTCGGCCCCCTCTTTCGCCCCACTGCTGT 1009  
 DB 181 CTGGGCCCCCAGGACGAGAGGCTGTCGGCCCCCTCTTTCGCCCCACTGCTGT 240  
 QY 1010 CCATGCGGGGCGCCGCTTGGCTTTCCGGAACCTGGGGGCCCTGCTTCCCGGCTGCACCG 1069  
 DB 241 CCATGCGGGGCGCCGCTTGGCTTTCCGGAACCTGGGGGCCCTGCTTCCCGGCTGCACCG 300  
 QY 1070 CTGTGCTGCGCGATGCGCCGACCCCTGCGCGGCTCTTCTGGCTGAGCTGTGACGCTGG 1129  
 DB 301 CTGTGCTGCGCGATGCGCCGACCCCTGCGCGGCTCTTCTGGCTGAGCTGTGACGCTGG 360  
 QY 1130 ATGGCACTCATGACTTTCAGCTGTTTACACGATTTCTGCGGCGGCTGTACCG 1189  
 DB 361 ATGGCACTCATGACTTTCAGCTGTTTACACGATTTCTGCGGCGGCTGTACCG 420  
 QY 1190 GCGCTGCCAGAGCTGAGCGGGGACCCGAGGCGCGGAGACACTATGATGAA 1240  
 DB 421 GCGCTGCCAGAGCTGAGCGGGGACCCGAGGCGCGGAGACACTATGATGAAAGGCTT 480  
 QY 1241 ----- 1240  
 DB 481 CTGGCTGCTAGGAGTCTGATCAGAGTCTGTCGCCAGTTTGACAGAAGGAGCGGA 540  
 QY 1241 ----- 1240  
 DB 541 GCTTATTCAAGTCTAGAGGAGTGGAGAGTTAAGCGCTGGATTTTCAGATCTCCCTGGTT 600  
 QY 1241 ----- 1240

DB 601 CCAGCCGCGAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATATCTCACCAGCGCCTT 660  
 QY 1241 ----- 1240  
 DB 661 CCAGCTCAGGCGTCTAGNAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCTC 720  
 QY 1241 ----- 1240  
 DB 721 TCACCCGCGTGTCTCAGAGCTGAGACTCCCAAGAAACCTTTCAGACTACCTTCTCTCTGCC 780  
 QY 1241 ----- 1240  
 DB 781 TTCAGCAAGGGCGTGTGCCACATTTCTGAGGGGCTTCGGATGGCCAGCTTGGGGCTGT 840  
 QY 1269 TCCTCAGTGGCCCATCTCCCTGGTCTCTCTCTGTCATGAGACCGGCTGGTGCAGCGAT 1328  
 DB 841 TCCTCAGTGGCCCATCTCCCTGGTCTCTCTCTGTCATGAGACCGGCTGGTGCAGCGAT 900  
 QY 1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGTGGCTGCCGTGCCA 1388  
 DB 901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGTGGCTGCCGTGCCA 960  
 QY 1389 CATGCTGTCCACACAGTGTGGCCGCTGACAGCTTTCAGCGCCCTCACCAGCTTCACCT 1448  
 DB 961 CATGCTGTCCACACAGTGTGGCCGCTGACAGCTTTCAGCGCCCTCACCAGCTTCACCT 1020  
 QY 1449 TCTCAGCCCTGCAGATCTCTGCCCTACACACTGGCCCTCCCTCTACCCACGGGAGAGCAG 1508  
 DB 1021 TCTCAGCCCTGCAGATCTCTGCCCTACACACTGGCCCTCCCTCTACCCACGGGAGAGCAG 1080  
 QY 1509 TGTTCCTGCCCAATACGAGGGGACACTGGAGTGTCTAGCAGTGAGGACACCTGTATGA 1568  
 DB 1081 TGTTCCTGCCCAATACGAGGGGACACTGGAGTGTCTAGCAGTGAGGACACCTGTATGA 1140  
 QY 1569 CCAGCTTCTCTGCGAGCCCTAAGCCTGGAGCTTCCCTTCCCTAAATGACACCTGGGTGCTG 1628  
 DB 1141 CCAGCTTCTCTGCGAGCCCTAAGCCTGGAGCTTCCCTTCCCTAAATGACACCTGGGTGCTG 1200  
 QY 1629 GAGGAGTGGCTGTCTCCACCTCCACCCGCGCTCTGCGGGGCTCTGCCCTGTGATGCT 1688  
 DB 1201 GAGGAGTGGCTGTCTCCACCTCCACCCGCGCTCTGCGGGGCTCTGCCCTGTGATGCT 1260  
 QY 1689 CCGTACGTGTGTGTGGTGGTGGCCACCGAGGCGAGGCTGTCCCGGGCGGGGCACT 1748  
 DB 1261 CCGTACGTGTGTGTGGTGGTGGCCACCGAGGCGAGGCTGTCCCGGGCGGGGCACT 1320  
 QY 1749 GCGTGGACCTCGCCATCTCGATAGTGTCTCTCTGTCAGGTGGCCGCTCCCTGTGATGCT 1808  
 DB 1321 GCGTGGACCTCGCCATCTCGATAGTGTCTCTCTGTCAGGTGGCCGCTCCCTGTGATGCT 1380  
 QY 1809 TTATGGGTCTCATTTGTCAGCTCAGCAGTGTCTGCTCCTATATGCTGTGTCGCGGAG 1868  
 DB 1381 TTATGGGTCTCATTTGTCAGCTCAGCAGTGTCTGCTCCTATATGCTGTGTCGCGGAG 1440  
 QY 1869 GCGTGGGTCTGCTGCCATTTACTTTGTACACAGTAGTAGTTTACACAGGAGGAGGCTGG 1928  
 DB 1441 GCGTGGGTCTGCTGCCATTTACTTTGTACACAGTAGTAGTTTACACAGGAGGAGGCTGG 1500  
 QY 1929 CCAATACTCAGCTAGAAAACCTTCCAGCACATTTGGGCTGGAGGCGCTGCCCTCACTGGGT 1988  
 DB 1501 CCAATACTCAGCTAGAAAACCTTCCAGCACATTTGGGCTGGAGGCGCTGCCCTCACTGGGT 1560  
 QY 1989 CCCAGCTCCCGCTCTGTAGCCCCATGGGCTGCCGGGCTGGCGGCGAGTTTCTGTTG 2048  
 DB 1561 CCCAGCTCCCGCTCTGTAGCCCCATGGGCTGCCGGGCTGGCGGCGAGTTTCTGTTG 1620  
 QY 2049 CTGCCAAAGTAAATGTGGTCTCTGCTGCCACCTCTGCTGTGAGGTGCGTAGCTGCACA 2108  
 DB 1621 CTGCCAAAGTAAATGTGGTCTCTGCTGCCACCTCTGCTGTGAGGTGCGTAGCTGCACA 1680  
 QY 2109 GCTGGGGGCTGGGGGCTCCCTCTCTCTCTCCCCAGTCTCTAGGGCTGCCCTGACTGGAGG 2168  
 DB 1681 GCTGGGGGCTGGGGGCTCCCTCTCTCTCTCCCCAGTCTCTAGGGCTGCCCTGACTGGAGG 1740



QY	2169	C	TTTCCAAAGGGGTTTTCAGTCTGACCTTATACAGGGAGGCCAGAGGGCTTCCATGCACCTG	2228	3247	G	GTGAATATTTTATACGTAGTGCAGCAATCAGAGTATATGTTTATGTCACAAATTA	3300
Db	1741	C	TTTCCAAAGGGGTTTTCAGTCTGACCTTATACAGGGAGGCCAGAGGGCTTCCATGCACCTG	1800	2821	G	GTGAATATTTTATACGTAGTGCAGCAATCAGAGTATATGTTTATGTCACAAATTA	2880
QY	2229	G	AATCGGGGACTCTGAGAGTGGATATCCAGAGCTCAGGGTTAACAGCTAGCCTCTCTAGT	2288	3307	A	AGGCTTTCTTATATGTTTAAAAA	3330
Db	1801	G	AATCGGGGACTCTGAGAGTGGATATCCAGAGCTCAGGGTTAACAGCTAGCCTCTCTAGT	1860	2881	A	AGGCTTTCTTATATGTTTAAAAA	2904
QY	2289	T	GACACACCTCAGAGAGGGTTTTTGGAGCTGAATAAACTCAGTCACCTGCTTCCCA	2348				
Db	1861	T	GACACACCTCAGAGAGGGTTTTTGGAGCTGAATAAACTCAGTCACCTGCTTCCCA	1920				
QY	2349	T	CTCTAAAGCCCTTAAACCTGCAGCTCGTTTAAATGTAGCTCTTGCATGGAGTTCTTAGG	2408				
Db	1921	T	CTCTAAAGCCCTTAAACCTGCAGCTCGTTTAAATGTAGCTCTTGCATGGAGTTCTTAGG	1980				
QY	2409	A	TGAACACTCTCCATGAGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCTCTG	2466				
Db	1981	A	TGAACACTCTCCATGAGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCTCTG	2040				
QY	2467	A	GGGGCAACACACAGAACACAGTCCCTCAGCCACAGCCACAGCACTGCTCTTTTGTGATCCA	2526				
Db	2041	A	GGGGCAACACACAGAACACAGTCCCTCAGCCACAGCCACAGCACTGCTCTTTTGTGATCCA	2100				
QY	2527	C	CCCCCTCTTACCTTTTATCAGGATGTGGCTGTGTGGTCTCTGTGTCATCACAGAGA	2586				
Db	2101	C	CCCCCTCTTACCTTTTATCAGGATGTGGCTGTGTGGTCTCTGTGTCATCACAGAGA	2160				
QY	2587	C	CACAGCATTTAAATTAATTAATTTATTTAAACAAAGTAGAAGGAATCCATTTGCTAG	2646				
Db	2161	C	CACAGCATTTAAATTAATTTATTTTAAACAAAGTAGAAGGAATCCATTTGCTAG	2220				
QY	2647	C	TTTCTGTGTGGTGTCTAATTTTGGTAGGGTGGGGATCCCAACAATCAGTCTCC	2706				
Db	2221	C	TTTCTGTGTGGTGTCTAATTTTGGTAGGGTGGGGATCCCAACAATCAGTCTCC	2280				
QY	2707	C	TGAGATAGCTGGTCTATTTGCGGCTGATCTTCCAGAAATCTTCTCTCTGGGTCTGGCC	2766				
Db	2281	C	TGAGATAGCTGGTCTATTTGCGGCTGATCTTCCAGAAATCTTCTCTCTGGGTCTGGCC	2340				
QY	2767	C	CCCAAAATGCCATAACCCAGGACCTTGAAATTTCTACTATCCCAAAATGATAATCCAAA	2826				
Db	2341	C	CCCAAAATGCCATAACCCAGGACCTTGAAATTTCTACTATCCCAAAATGATAATCCAAA	2400				
QY	2827	T	GCTGTTACCAAGGTTAGGTTTGAAGGAGGTAGAGGGTGGGGCTTCAAGTCTCAAC	2886				
Db	2401	T	GCTGTTACCAAGGTTAGGTTTGAAGGAGGTAGAGGGTGGGGCTTCAAGTCTCAAC	2460				
QY	2887	G	CTTCCCTTAACACCCCTCTTCTCTTGGCCAGCGCTGTTTCCCTCCCTTCCACTCCCC	2946				
Db	2461	G	CTTCCCTTAACACCCCTCTTCTCTTGGCCAGCGCTGTTTCCCTCCCTTCCACTCCCC	2520				
QY	2947	T	CTACTCTCTTAGGACTGGGTGATGAAGCACTGCCCCAAAATTTCCCTTACCCCCAAC	3006				
Db	2521	T	CTACTCTCTTAGGACTGGGTGATGAAGCACTGCCCCAAAATTTCCCTTACCCCCAAC	2580				
QY	3007	T	TTTCCCTTACCCCAACTTTCCCAAGCTTCCCAACCCCTGTTTGGAGCTACTGCAGGA	3066				
Db	2581	T	TTTCCCTTACCCCAACTTTCCCAAGCTTCCCAACCCCTGTTTGGAGCTACTGCAGGA	2640				
QY	3067	C	CAGAGCAAAAGTGGGTTTCCCAAGCTTTTGTCCATCTCAGCCCCCAGATATATCT	3126				
Db	2641	C	CAGAGCAAAAGTGGGTTTCCCAAGCTTTTGTCCATCTCAGCCCCCAGATATATCT	2700				
QY	3127	G	TGCTTTGGGGAATCTCACACAGAACTCAGGAGCACCCCTTGCCTGAGCTAAGGAGGTC	3186				
Db	2701	G	TGCTTTGGGGAATCTCACACAGAACTCAGGAGCACCCCTTGCCTGAGCTAAGGAGGTC	2760				
QY	3187	T	TATCTCTCAGGGGGGTTTAAAGTGCCTTTTGAATAATATGTCGCTTATTTATTAGGGG	3246				
Db	2761	T	TATCTCTCAGGGGGGTTTAAAGTGCCTTTTGAATAATATGTCGCTTATTTATTAGCGG	2820				

QY	3247	G	GTGAATATTTTATACGTAGTGCAGCAATCAGAGTATATGTTTATGTCACAAATTA	3300
Db	2821	G	GTGAATATTTTATACGTAGTGCAGCAATCAGAGTATATGTTTATGTCACAAATTA	2880
QY	3307	A	AGGCTTTCTTATATGTTTAAAAA	3330
Db	2881	A	AGGCTTTCTTATATGTTTAAAAA	2904

RESULT 20			
AS64038	ID AAS64038 standard; cDNA; 4894 BP.		
XX	AC	AAS64038;	
XX	AC	AAS64038;	
DT	29-JAN-2002	(first entry)	
DE	Human prostate cancer; ss; cytostatic; immunostimulant; tumour.		
DE	Human prostate cancer; ss; cytostatic; immunostimulant; tumour.		
XX	Homo sapiens.		
XX	WO200173032-A2.		
XX	04-OCT-2001.		
XX	27-MAR-2001; 2001WO-US09919.		
XX	27-MAR-2000; 2000US-0536857.		
XX	09-MAY-2000; 2000US-0568100.		
XX	12-MAY-2000; 2000US-0570737.		
XX	13-JUN-2000; 2000US-0593793.		
XX	27-JUN-2000; 2000US-0605783.		
XX	10-AUG-2000; 2000US-0636215.		
XX	29-AUG-2000		





Db 3471 ATTTGACAGAGCGACTTGGCCAAATACCTACCGTAGAATACTCCAGCACAATTGGGGTG 3530  
 QY 1969 GAGGGCTCCCTCAGCTGAGGCTCCAGCTCCCGCTCTCTGTTAGCCCATGGGGCTCCGGG 2028  
 Db 3531 GAGGGCTCCCTCAGCTGAGGCTCCAGCTCCCGCTCTCTGTTAGCCCATGGGGCTCCGGG 3590  
 QY 2029 CTGGCCGCCAGTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGTCGACACCTGTGCTG 2088  
 Db 3591 CTGGCCGCCAGTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGTCGACACCTGTGCTG 3650  
 QY 2089 CTGAGTGGCTGAGTCAGCAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 2148  
 Db 3651 CTGAGTGGCTGAGTCAGCAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTG 3710  
 QY 2149 TAGGGCTGCTGACTGGAGGCTTCCAGGGGGTTCAGCTGAGCTTATACAGGAGGC 2208  
 Db 3711 TAGGGCTGCTGACTGGAGGCTTCCAGGGGGTTCAGCTGAGCTTATACAGGAGGC 3770  
 QY 2209 CAGAAGGGCTCCATGACCTGGAATGCGGGACTCTCGAGGTGGATTACCCAGGCTCAGGG 2268  
 Db 3771 CAGAAGGGCTCCATGACCTGGAATGCGGGACTCTCGAGGTGGATTACCCAGGCTCAGGG 3830  
 QY 2269 TTAACAGCTAGCTTCTAGTGTGAGACACACCTAGAGAGGGTTTGGGAGCTGAATAA 2328  
 Db 3831 TTAACAGCTAGCTTCTAGTGTGAGACACACCTAGAGAGGGTTTGGGAGCTGAATAA 3890  
 QY 2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCTTAACTGACGCTTCTGTTTAACTAGCT 2388  
 Db 3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCTTAACTGACGCTTCTGTTTAACTAGCT 3950  
 QY 2389 CTTCATGGGAGTTTCTAGGATGAACACTCTCTCCATGGATTGAACATATG--ACTTA 2446  
 Db 3951 CTTCATGGGAGTTTCTAGGATGAACACTCTCCATGGATTGAACATATG--ACTTA 4010  
 QY 2447 TTGTAGGGAAGAGTCTGTAGGGGCAACACAGAACCCAGGTCCCTCAGCCACAGC 2506  
 Db 4011 TTGTAGGGAAGAGTCTGTAGGGGCAACACAGAACCCAGGTCCCTCAGCCACAGC 4070  
 QY 2507 ACTGCTCTTTGCTGATCAGCCCTCTTACCTTTATCAGGATGCGCTGTTGGTCC 2566  
 Db 4071 ACTGCTCTTTGCTGATCAGCCCTCTTACCTTTATCAGGATGCGCTGTTGGTCC 4130  
 QY 2567 TTCTGTTGGCATCACAGACACAGGCAATTAATATTTAACTTATTTAAACAAAGT 2626  
 Db 4131 TTCTGTTGGCATCACAGACACAGGCAATTAATATTTAACTTATTTAAACAAAGT 4190  
 QY 2627 AGAAGGAATCCATGCTAGCTTTCTGTGGTGTCTAATATTTGGTAGGGTGGGG 2686  
 Db 4191 AGAAGGAATCCATGCTAGCTTTCTGTGGTGTCTAATATTTGGTAGGGTGGGG 4250  
 QY 2687 ATCCCAACAAATCAGCTCCCTGAGATAGCTGCTATTTGGGCTGATCTGCAGAACT 2746  
 Db 4251 ATCCCAACAAATCAGCTCCCTGAGATAGCTGCTATTTGGGCTGATCTGCAGAACT 4310  
 QY 2747 TCTTCTCTGGGCTCTGGCCCCCAAAATGCCCTTAACCCAGGACTTGGAAATCTACTCA 2806  
 Db 4311 TCTTCTCTGGGCTCTGGCCCCCAAAATGCCCTTAACCCAGGACTTGGAAATCTACTCA 4370  
 QY 2807 TCCCAATGATAATCCAAATGCTTGTACCAAGTTAGGGTGTGAAGGAAGGTAGAG 2866  
 Db 4371 TCCCAATGATAATCCAAATGCTTGTACCAAGTTAGGGTGTGAAGGAAGGTAGAG 4430  
 QY 2867 GTGGGCTTCAGTCTCAGGCTTCCCTAACCCACCCCTCTCTCTTGGCCAGCTGTG 2926  
 Db 4431 GTGGGCTTCAGTCTCAGGCTTCCCTAACCCACCCCTCTCTCTTGGCCAGCTGTG 4490  
 QY 2927 TCCCCCACTTCCACTCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGACTGCCCA 2986  
 Db 4491 TCCCCCACTTCCACTCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGACTGCCCA 4550  
 QY 2987 AAATTTTCCCTACCCCAACTTTTCCCTACCCCAACTTTTCCCAACAGCTCCCAACCC 3046  
 Db 4551 AAATTTTCCCTACCCCAACTTTTCCCTACCCCAACTTTTCCCAACAGCTCCCAACCC 4610

QY 3047 TGTGTTGGAGCTACTGACAGGACCAAGACACAAAGTGGGTTTCCCAAGCCTTTGTCCATC 3106  
 Db 4611 TGTGTTGGAGCTACTGACAGGACCAAGACACAAAGTGGGTTTCCCAAGCCTTTGTCCATC 4670  
 QY 3107 TCAGCCCCCAGAGTATATCTGTGTTGGGGAATCTACACAGAAACTCAGAGACACCCC 3166  
 Db 4671 TCAGCCCCCAGAGTATATCTGTGTTGGGGAATCTACACAGAAACTCAGAGACACCCC 4730  
 QY 3167 TGCCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAAGTGGCGTTTGCATTAATG 3226  
 Db 4731 TGCCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAAGTGGCGTTTGCATTAATG 4790  
 QY 3227 TCGTCTTATTTATTTAGCGGGTCAATATTTTATCTGTAAGTCAGCAATCAGAGTATAA 3286  
 Db 4791 TCGTCTTATTTATTTAGCGGGTCAATATTTTATCTGTAAGTCAGCAATCAGAGTATAA 4850  
 QY 3287 TGTGTTATGGTGACAAAATTTAAAGCTTCTTATATGTTTAAAAA 3330  
 Db 4851 TGTGTTATGGTGACAAAATTTAAAGCTTCTTATATGTTTAAAAA 4894

RESULT 21  
 AAH93866  
 ID AAH93866 standard; cDNA; 4894 BP.  
 AC AAH93866;  
 XX 04-OCT-2001 (first entry)  
 DE P553S cDNA splice variant P553S-14.  
 KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
 KW cytostatic; gene therapy; metastasis; ss.  
 OS Homo sapiens.  
 XX WO200151633-A2.  
 PN 19-JUL-2001.  
 PD 16-JAN-2001; 2001WO-US01574.  
 PF 14-JAN-2000; 2000US-0483672.  
 PR (CORI-) CORIXA CORP.  
 PA Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 PI Kallos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;  
 PI Wang A, Meagher MJ;  
 XX WPI: 2001-425873/45.  
 DR New polynucleotide encoding a prostate-specific protein, for  
 PT diagnosing, monitoring and treating prostate cancer in a patient and  
 PS for use in vaccines  
 CC Claim 1; Page 457-459; 543pp; English.  
 XX The present invention describes polynucleotide sequences (I) which encode  
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
 CC and can be used in vaccine production and gene therapy. (I), (II),  
 CC antibodies to (II), fusion proteins comprising (II), and isolated  
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.  
 CC (I) and the antibodies are also used in the detection of cancer in a  
 CC patient. The cancer that is diagnosed or treated is particularly  
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
 CC (I) can be used for monitoring the progression of cancer in a patient.  
 CC (I) and (II) can also be used to improve diagnostic and therapeutic  
 CC methods for prostate cancer. They can indicate the level of metastasis  
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to  
 CC AAH01318 represent polynucleotide and amino acid sequences used in the

CC exemplification of the present invention.

```
XX Sequence 4894 BP; 928 A; 1448 C; 1354 G; 1163 T; 1 other;
SQ

Query Match      62.8%; Score 2142.8; DB 22; Length 4894;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 2872; Conservative 1; Mismatches 3; Indels 708; Gaps 2;

QY 455 GGCATTGGTCCAGTGTGGGCTGTGTCGTGTCCTGCTAGGCTCAGCCAGTACCAC 514
    |||||
Db 1311 GGCATTGGTCCAGTGTGGGCTGTGTCGTGTCCTGCTAGGCTCAGCCAGTACCAC 1370

QY 515 TGGCGTGGAGCTATGGCGCGCGGCTTCACTGGGCTGCTGGGCTGCTGGGCTG 574
    |||||
Db 1371 TGGCGTGGAGCTATGGCGCGCGGCTTCACTGGGCTGCTGGGCTGCTGGGCTG 1430

QY 575 CTGAGGCTCTTTCTATCCCAAGGGCGGCTGGCTAGCAGGGCTGCTGCGCGGATGCC 634
    |||||
Db 1431 CTGAGGCTCTTTCTATCCCAAGGGCGGCTGGCTAGCAGGGCTGCTGCGCGGATGCC 1490

QY 635 AGCCGCTGGAGTGGCACTGCTCATCTGGGCGTGGGCTGCTGGACTTCTGTGGCGAG 694
    |||||
Db 1491 AGCCGCTGGAGTGGCACTGCTCATCTGGGCGTGGGCTGCTGGACTTCTGTGGCGAG 1550

QY 695 GTGTGCTTCACTCCACTGGAGGCTGCTGCTCTGACCTTCTCCGGGACCGGACACTGT 754
    |||||
Db 1551 GTGTGCTTCACTCCACTGGAGGCTGCTGCTCTGACCTTCTCCGGGACCGGACACTGT 1610

QY 755 CGCCAGGCTACTCTGTCTATGCTTCTATGATGATGATGATGATGATGATGATGATG 814
    |||||
Db 1611 CGCCAGGCTACTCTGTCTATGCTTCTATGATGATGATGATGATGATGATGATGATG 1670

QY 815 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874
    |||||
Db 1671 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1730

QY 875 TGCTCTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 934
    |||||
Db 1731 TGCTCTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1790

QY 935 GCTGAGGAGGAGCGGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 994
    |||||
Db 1791 GCTGAGGAGGAGCGGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1850

QY 995 TCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
    |||||
Db 1851 TCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1910

QY 1055 CCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1114
    |||||
Db 1911 CCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1970

QY 1115 GAGCTGTGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1174
    |||||
Db 1971 GAGCTGTGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2030

QY 1175 GAGGCTGTACAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234
    |||||
Db 2031 GAGGCTGTACAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2090

QY 1235 GATGA----- 1239
    |||||
Db 2091 GATGAAGGTAAAGGCTTGGCAGCCAGCAGAGGCTGGTGGGAGCGCGCCACAGAGAG 2150

QY 1240 ----- 1239
    |||||
Db 2151 ACACCTGGGGCTGTGTCTGGGCTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2210

QY 1240 ----- 1239
    |||||
Db 2211 AAGTGGGGATGGAGCCCATGTGCATACAGGCTTCTCATGGGTGTGGAACATCTGTCT 2270

QY 1240 ----- 1239
    |||||
```

3351	OCAGGTGCCCATCCCTCTGTTTATATGGGCTCAATGTTCAGCTCAGCCAGTCTGTCACTGC	3410
1849	CTATATGGTGTCTGCCGAGGCCCTGGGTCTGGTCGCCATTTACTTTTGCTACACAGGTAGT	1908
3411	CTATATGGTGTCTGCCGAGGCCCTGGGTCTGGTCGCCATTTACTTTTGCTACACAGGTAGT	3470
1909	ATTTGACAAGAGCGACTTGTGGCCAAATACTACGCGTAGAAAACTTCCAGCACATTTGGGGTG	1968
3471	ATTTGACAAGAGCGACTTGGCCAAATACTACGCGTAGAAAACTTCCAGCACATTTGGGGTG	3530
1969	GAGGCCCTGCCTCACTGGGTCCCAGCTCCCGCTCCCTGTAGGCCCATGGGGCTGCCGGG	2028
3531	GAGGCCCTGCCTCACTGGGTCCCAGCTCCCGCTCCCTGTAGGCCCATGGGGCTGCCGGG	3590
2029	CTGGCCGCCAGTTTCTGTGTCTGTCGCAAAAGTAATGTGGCTCTCTGCTGCCACCCCTGTGCTG	2088
3591	CTGGCCGCCAGTTTCTGTGTCTGTCGCAAAAGTAATGTGGCTCTCTGCTGCCACCCCTGTGCTG	3650
2089	CTGAGGTGCGTAGCTGACAGCTGGGGGCTGGGGGCTCCCTCTCTCTCTCCCGAGTCTC	2148
3651	CTGAGGTGCGTAGCTGACAGCTGGGGGCTGGGGGCTCCCTCTCTCTCTCCCGAGTCTC	3710
2149	TAGGGCTGCCCTCACTGGAGGCTTCCAGAGGGGTTTTCAGTCTGGACTTTATACAGGAGGC	2208
3711	TAGGGCTGCCCTCACTGGAGGCTTCCAGAGGGGTTTTCAGTCTGGACTTTATACAGGAGGC	3770
2209	CAGAAGGCTCCATGCATCGAATCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG	2268
3771	CAGAAGGCTCCATGCATCGAATCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG	3830
2269	TTAACAGCTAGCCTCCTTAGTTGAGACACACCTTAGAGAAGGGTTTTTGGGAGCTGAATAAA	2328
3831	TTAACAGCTAGCCTCCTTAGTTGAGACACACCTTAGAGAAGGGTTTTTGGGAGCTGAATAAA	3890
2329	CTCAGTACCTGGTTTTCCCATCTCTAAGCCCTTAACTGCGAGCTTCGTTAATCTAGCT	2388
3891	CTCAGTACCTGGTTTTCCCATCTCTAAGCCCTTAACTGCGAGCTTCGTTAATCTAGCT	3950
2389	CTTGATGGGAGTTTCTAGGATGAACACTCCTCCATGGGATTTGAAACATATG - ACTTA	2446
3951	CTTGATGGGAGTTTCTAGGATGAACACTCCTCCATGGGATTTGAAACATATGAAAGTTA	4010
2447	TTTGTAGGGGAAGACTCTGAGGGGCAACACAAAGACCAGGTCCCTCAGCCCAACAGC	2506
4011	TTTGTAGGGGAAGACTCTGAGGGGCAACACAAAGACCAGGTCCCTCAGCCCAACAGC	4070
2507	ACTGTCTTTTGGCTGATCACCCCTCTTACTTTTATCAGSAGTGTGGCTGTGGTCC	2566
4071	ACTGTCTTTTGGCTGATCACCCCTCTTACTTTTATCAGSAGTGTGGCTGTGGTCC	4130
2567	TTCTGTTCGCATCACAGACACAGGCATTTAAATATTAACTTATTTTAAACAAAGT	2626
4131	TTCTGTTCGCATCACAGACACAGGCATTTAAATATTAACTTATTTTAAACAAAGT	4190
2627	AGAAGGGAATCCATGCTAGCTTTCTGTGTGGTGCTCTAAATATTTGGGTAGGTGGGG	2686
4191	AGAAGGGAATCCATGCTAGCTTTCTGTGTGGTGCTCTAAATATTTGGGTAGGTGGGG	4250
2687	ATCCCCAACAACTCAGTCCCTCAGATAGCTGGTCAATTGGGCTGATCATTTGCCAGAATCT	2746
4251	ATCCCCAACAACTCAGTCCCTCAGATAGCTGGTCAATTGGGCTGATCATTTGCCAGAATCT	4310
2747	TCCTCTCTGGGTCTGGCCCCCAAAATGCCCTAACCCAGGACCTTTGGAAATCTACTCA	2806
4311	TCCTCTCTGGGTCTGGCCCCCAAAATGCCCTAACCCAGGACCTTTGGAAATCTACTCA	4370
2807	TCCCAATGATAATTCCAAATGCTGTTACCCAAAGTTAGGGTGTGAAGGAAGGTAGAG	2866
4371	TCCCAATGATAATTCCAAATGCTGTTACCCAAAGTTAGGGTGTGAAGGAAGGTAGAG	4430
2867	GTGGGGCTTCAGGTCTCAAGGGCTTCCTTAACACACCCCTCTCTCTTGGCCACGCTGGT	2926
4431	GTGGGGCTTCAGGTCTCAAGGGCTTCCTTAACACACCCCTCTCTCTTGGCCACGCTGGT	4490

QY	2927	TCCTCCCTACCTTCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCCA	2986
Db	4491	TCCTCCCTACCTTCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCCA	4550
QY	2987	AAATTTCCCTACCTCCCTACCTTTCCCTCTACCTCCCTACCTTTCCCTACCTCCCAACCC	3046
Db	4551	AAATTTCCCTACCTCCCTACCTTTCCCTCTACCTCCCTACCTTTCCCTACCTCCCAACCC	4610
QY	3047	TGTTTGAGAGCTACTGCAGGACCAAGAGCACAAGTCCGGTTTCCCAAGCTTTGTCCATC	3106
Db	4611	TGTTTGAGAGCTACTGCAGGACCAAGAGCACAAGTCCGGTTTCCCAAGCTTTGTCCATC	4670
QY	3107	TCAGCCCCCAGAGTATATCTGTCTTGGGGAATCTCACACAGAACTCAGGAGCACCCTCC	3166
Db	4671	TCAGCCCCCAGAGTATATCTGTCTTGGGGAATCTCACACAGAACTCAGGAGCACCCTCC	4730
QY	3167	TGCTGAGCTAAGGGAGGCTCTTATCTCTCAGGGGGGTTTAACTGCGCTTTGCAATAATG	3226
Db	4731	TGCTGAGCTAAGGGAGGCTCTTATCTCTCAGGGGGGTTTAACTGCGCTTTGCAATAATG	4790
QY	3227	TCGCTTATTTATTTAGCGGGGTGAATATTTTATCTGTAAGTGAGCAATCAGAGTATAA	3286
Db	4791	TCGCTTATTTATTTAGCGGGGTGAATATTTTATCTGTAAGTGAGCAATCAGAGTATAA	4850
QY	3287	TGTTTATGCTGACAAAATTAAGGCTTCTTATATGTTTAAAAA	3330
Db	4851	TGTTTATGCTGACAAAATTAAGGCTTCTTATATGTTTAAAAA	4894
RESULT 22			
ABL95409			
ID	ABL95409 standard; cDNA; 4894 BP.		
XX	ABL95409;		
XX	19-JUL-2002 (first entry)		
XX	Human P53S splice variant SEQ ID NO 702.		
DE	Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;		
KW	gene therapy; gene; ss.		
OS	Homo sapiens.		
XX	US2002022248-A1.		
PN	21-FEB-2002.		
PD	12-JAN-2001; 2001US-0759143.		
PF	25-FEB-1997; 97US-0806099.		
XX	01-AUG-1997; 97US-0904804.		
PR	09-FEB-1998; 98US-0020956.		
PR	25-FEB-1998; 98US-0030607.		
PR	14-JUL-1998; 98US-0115453.		
PR	23-SEP-1998; 98US-0159812.		
PR	15-JAN-1999; 98US-0232149.		
PR	09-APR-1999; 99US-0288946.		
PR	13-JUL-1999; 99US-0352616.		
PR	12-NOV-1999; 99US-0439313.		
PR	18-NOV-1999; 99US-0443686.		
PR	14-JAN-2000; 2000US-0483672.		
PR	27-MAR-2000; 2000US-0536857.		
PR	09-MAY-2000; 2000US-0568100.		
PR	12-MAY-2000; 2000US-0570737.		
PR	13-JUN-2000; 2000US-0593793.		
PR	27-JUN-2000; 2000US-0605783.		
PR	10-AUG-2000; 2000US-0636215.		
PR	29-AUG-2000; 2000US-0651236.		
PR	06-SEP-2000; 2000US-0651279.		
PR	02-OCT-2000; 2000US-0679426.		
PR	10-OCT-2000; 2000US-0685166.		

XX (XUJ/J) XU J.  
 PA (DILL/) DILLON D C.  
 PA (MITC/) MITCHAM J L.  
 PA (HARL/) HARLOCKER S L.  
 PA (JIAN/) JIANG Y.  
 PA (KALO/) KALOS M D.  
 PA (FANG/) FANGER G R.  
 PA (RETT/) RETTER M W.  
 PA (STOL/) STOLK J A.  
 PA (DAYC/) DAY C H.  
 PA (VEDV/) VEDWICK T S.  
 PA (CART/) CARTER D.  
 PA (LISX/) LI S X.  
 PA (WANG/) WANG A.  
 PA (SKEI/) SKEIKY Y A W.  
 PA (HEPL/) HEPLER W T.  
 PA (HEND/) HENDERSON R A.  
 XX  
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;  
 XX  
 XX WPI: 2002-255649/30.  
 XX  
 XX New prostate-specific polynucleotides for diagnosing and treating  
 PT diseases, in particular prostate cancer, and as markers for the  
 PT progression of cancer  
 XX  
 XX Claim 1; SEQ ID NO 702; 87pp; English.  
 PS  
 XX  
 CC The present invention provides prostate-specific coding sequences and  
 CC their encoded proteins. These can be used in the diagnosis and treatment  
 CC of cancers, particularly prostate cancer. The present sequence is a cDNA  
 CC described in the invention.  
 XX  
 SQ Sequence 4894 BP; 928 A; 1448 C; 1354 G; 1163 T; 1 other;  
 Query Match 62.8%; Score 2142.8; DB 24; Length 4894;  
 Best Local Similarity 80.1%; Pred. No. 0;  
 Matches 2872; Conservative 1; Mismatches 3; Indels 708; Gaps 2;  
 455 GCATTGGTCCAGTGTGGGCTGTGTGTGTCGCGCTCTAGGCTCAGCCAGTACCAC 514  
 1311 GGCATTGGTCCAGTGTGGGCTGTGTGTGTCGCGCTCTAGGCTCAGCCAGTACCAC 1370  
 515 TGGCGTGGACGATATGGCGCGCGCGCGCTTCATCTGGGCACGTCTTGGGCATCTG 574  
 1371 TGGCGTGGACGATATGGCGCGCGCGCGCTTCATCTGGGCACGTCTTGGGCATCTG 1430  
 575 CTGAGCGCTCTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGCTGTGCCCGATCCC 634  
 1431 CTGAGCGCTCTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGCTGTGCCCGATCCC 1490  
 635 AGGCCCTGGAGCTGGCAGTGTCTATCTGCGCGTGGGCTGCTGGACTTCTGTGGCAG 694  
 1491 AGGCCCTGGAGCTGGCAGTGTCTATCTGCGCGTGGGCTGCTGGACTTCTGTGGCAG 1550  
 695 GTGTGCTTCACTCCACTGGAGCGCTGCTCTGACCTTCTCCGGGACCGGACCACTGT 754  
 1551 GTGTGCTTCACTCCACTGGAGCGCTGCTCTGACCTTCTCCGGGACCGGACCACTGT 1610  
 755 CGCCAGGCGCTACTGTCTATGCTTCTATGATCATGATCTTGGGGCTGCTGGGCTACCTC 814  
 1611 CGCCAGGCGCTACTGTCTATGCTTCTATGATCATGATCTTGGGGCTGCTGGGCTACCTC 1670  
 815 CTGCGCTGCANTGACTGGGACACCAAGTCCCTGGCCCGCTTACCTGGGACCGAGGAG 874  
 1671 CTGCGCTGCANTGACTGGGACACCAAGTCCCTGGCCCGCTTACCTGGGACCGAGGAG 1730  
 875 TGCCCTTTGGCGCTGCTCACCCTCATCTTCTCCTACCTGCGTACGAGCCACACAGTCTGGTG 934  
 1731 TGCCCTTTGGCGCTGCTCACCCTCATCTTCTCCTACCTGCGTACGAGCCACACAGTCTGGTG 1790

QY 935 GCTGAGGAGGAGCGCTGGGCCCCCACCAGCAGCAGAGGGCTGTGCGGCCCTCCTTG 994  
 DB 1791 GCTGAGGAGGAGCGCTGGGCCCCCACCAGCAGCAGAGGGCTGTGCGGCCCTCCTTG 1850  
 QY 995 TCGCCCCACTGCTGTCCATATGCGGGCCCGCTTGGCTTTCCGGAACCTTGGCGCCCTGCTT 1054  
 DB 1851 TCGCCCCACTGCTGTCCATATGCGGGCCCGCTTGGCTTTCCGGAACCTTGGCGCCCTGCTT 1910  
 QY 1055 CCCGGCTGACACGCTGTGCTGGCGCATGCGCGCACCCCTGCGCGGCTTCTGCTGGCT 1114  
 DB 1911 CCCGGCTGACACGCTGTGCTGGCGCATGCGCGCACCCCTGCGCGGCTTCTGCTGGCT 1970  
 QY 1115 GAGCTGTGACGCTGGATGGCACTCATGACCTTACGCTGTGTTTACAGGATTTCTGCTGGC 1174  
 DB 1971 GAGCTGTGACGCTGGATGGCACTCATGACCTTACGCTGTGTTTACAGGATTTCTGCTGGC 2030  
 QY 1175 GAGGGCTGTACAGGGCGTCCAGAGCTGAGCGCGGACCCGAGGCGCGGAGACACTAT 1234  
 DB 2031 GAGGGCTGTACAGGGCGTCCAGAGCTGAGCGCGGACCCGAGGCGCGGAGACACTAT 2090  
 QY 1235 GATGA----- 1239  
 DB 2091 GATGAAGTAAGGCTTGGCAGCCAGCAGAGGCTGGTGTGGAGCCGCCACCAGAGACG 2150  
 QY 1240 ----- 1239  
 DB 2151 ACACCTCGGGCTGTGCTGGGCTGGTCCCTCTCCATCTGCCCCGACTTCTCTGTCAGG 2210  
 QY 1240 ----- 1239  
 DB 2211 AAAGTGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270  
 QY 1240 ----- 1239  
 DB 2271 TGGGTTTTCAGAGAGGCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCTGTGCCCAAGTT 2330  
 QY 1240 ----- 1239  
 DB 2331 TGACAGAAGAAAGCGGAGCTTATTCAAAGTCTAGAGGAGTGGAGGAGTTAAGGCTGG 2390  
 QY 1240 ----- 1239  
 DB 2391 ATTTTCAGATCTGCTGTTTCCAGCGCGAGTGTGCCCTCTGCTCCCCCAACGACTTTTCCAA 2450  
 QY 1240 ----- 1239  
 DB 2451 ATAATCTCACCAGCGCTTCCAGCTCAGGCGTCTTAGAAGCGTCTTGAAGCCTATGGCCA 2510  
 QY 1240 ----- 1239  
 DB 2511 GCTGTCTTTGTGTCCCTCTCACCCTGCTGCTCAGCTGAGACTCCAGGAACCTT 2570  
 QY 1240 ----- 1239  
 DB 2571 CAGACTACCTTCTCTGCTTTCAGCAAGGGCGCTTGCACACATTTCTGAGGGTCAAGTG 2630  
 QY 1240 ----- 1239  
 DB 2631 AAGAACTAGACTCCCATTTGCTAGAGGTAGAAAGGGAAGGTTGTTGGGAGCAGGCGCTG 2690  
 QY 1240 ----- 1239  
 DB 2691 GTCCACAGCAGGTCTCGTGAGCAGGATACCTGTGTGTTCCGCTTCTCATCTCCCTGAGAC 2750  
 QY 1240 -----AGGCTTCG 1248  
 DB 2751 TGCTCCGACCTTCCCTCCAGGCTCTGTGTGATGGCCCTCTCCCTCTGCGAGGCGTTCG 2810  
 QY 1249 GATGGCAGCGCTGGGCTGTTCCTGCAAGTGCAGTCCCTCCCTGCTCTCTCTGCTCAT 1308  
 DB 2811 GATGGCAGCGCTGGGCTGTTCCTGCAAGTGCAGTCCCTCCCTGCTCTCTCTGCTCAT 2870

QY 1309 GGACCGGCTGGTGCAGGATTCGGCAGCTCGAGCAGTCTATTGGCCAGTGTGCGAGCTTT 1368  
 DB 2871 GGACCGGCTGGTGCAGGATTCGGCAGCTCGAGCAGTCTATTGGCCAGTGTGCGAGCTTT 2930  
 QY 1369 CCCTGTGGCTGCCGGTGCCACATGCTCTGTCGCCAGTGTGCGCGTGGTGACAGCTTCAGC 1428  
 DB 2931 CCCTGTGGCTGCCGGTGCCACATGCTCTGTCGCCAGTGTGCGCGTGGTGACAGCTTCAGC 2990  
 QY 1429 CGCCCTCACCGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCCTCCCT 1488  
 DB 2991 CGCCCTCACCGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCCTCCCT 3050  
 QY 1489 CTACACCGGAGAGAGAGTGTCTCTGCGCCAAATACCGAGGGGACACTGGAGTGTCTAG 1548  
 DB 3051 CTACACCGGAGAGAGAGTGTCTCTGCGCCAAATACCGAGGGGACACTGGAGTGTCTAG 3110  
 QY 1549 CAGTGAGACAGCTGATGACAGCTTCTGCGCCAGGCGCTTAAGCCTGGAGCTCCCTTCCC 1608  
 DB 3111 CAGTGAGACAGCTGATGACAGCTTCTGCGCCAGGCGCTTAAGCCTGGAGCTCCCTTCCC 3170  
 QY 1609 TAATGGACAGCTGGTGTGAGGAGTGGCTGCTCCACCTCCACCGCGCTCTCGG 1668  
 DB 3171 TAATGGACAGCTGGTGTGAGGAGTGGCTGCTCCACCTCCACCGCGCTCTCGG 3230  
 QY 1669 GGCCTCTGCTGTGATGCTCCGTACGCTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1728  
 DB 3231 GGCCTCTGCTGTGATGCTCCGTACGCTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3290  
 QY 1729 GGTTCGGGCGGGGCACTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1788  
 DB 3291 GGTTCGGGCGGGGCACTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3350  
 QY 1789 CCAGTGGCGCCATCCCTGTTTATGGCTCCATGTCAGCTGTCAGCTGTCAGCTGTCAGCTG 1848  
 DB 3351 CCAGTGGCGCCATCCCTGTTTATGGCTCCATGTCAGCTGTCAGCTGTCAGCTGTCAGCTG 3410  
 QY 1849 CTATATGCTGTGCGCAGGCTGGTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1908  
 DB 3411 CTATATGCTGTGCGCAGGCTGGTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3470  
 QY 1909 ATTTGACAGAGCGATTTGGCAGTAATCTCAGCTGAGAAATCTCCAGCAGATTTGGGGTG 1968  
 DB 3471 ATTTGACAGAGCGATTTGGCAGTAATCTCAGCTGAGAAATCTCCAGCAGATTTGGGGTG 3530  
 QY 1969 GAGGCGCTGCTCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2028  
 DB 3531 GAGGCGCTGCTCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3590  
 QY 2029 CTGGCGCGCAGTTTCTGTTGCTGCGCAAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2088  
 DB 3591 CTGGCGCGCAGTTTCTGTTGCTGCGCAAGTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3650  
 QY 2089 CTGAGGTGGTGTGACAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2148  
 DB 3651 CTGAGGTGGTGTGACAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 3710  
 QY 2149 TAGGGCTGCTGACTGGAGGCTTCCAGGGGTTTCACTGTGAGTGTATACAGGAGGCT 2208  
 DB 3711 TAGGGCTGCTGACTGGAGGCTTCCAGGGGTTTCACTGTGAGTGTATACAGGAGGCT 3770  
 QY 2209 CAGAAGGCTCCATGCTGGAATGCGGGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAG 2268  
 DB 3771 CAGAAGGCTCCATGCTGGAATGCGGGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAG 3830  
 QY 2269 TTAAACAGTACCTCTAGTTGAGACACACTAGAGAGGGTGGTGGAGCTGTAATAA 2328  
 DB 3831 TTAAACAGTACCTCTAGTTGAGACACACTAGAGAGGGTGGTGGAGCTGTAATAA 3890  
 QY 2329 CTCAGTACCTGCTGCTTCCATCTCTAAGCCCTTAACTGTCAGCTTCTGTTAATGTAGCT 2388  
 DB 3891 CTCAGTACCTGCTGCTTCCATCTCTAAGCCCTTAACTGTCAGCTTCTGTTAATGTAGCT 3950  
 QY 2389 CTTGTCATGGGAGTTTCTAGGATGAAACACTCTCCATGGGATTTGAACATATG--ACTTA 2446

DB 3951 CTTGTCATGGAGCTTTCTAGGATGAAACACTCCACCATGGGATTTCAACATATGAAGTTA 4010  
 QY 2447 TTTGTAGGGAGAGTCTCTGAGGGCAACACACAGAACACAGTCCCTCAGCCACAGC 2506  
 DB 4011 TTTGTAGGGAGAGTCTCTGAGGGCAACACACAGAACACAGTCCCTCAGCCACAGC 4070  
 QY 2507 ACTGTCTTTTGTGATGATCCACCCCTCTTACCTTTTATCAGGATGTGCTGTTGGTCC 2566  
 DB 4071 ACTGTCTTTTGTGATGATCCACCCCTCTTACCTTTTATCAGGATGTGCTGTTGGTCC 4130  
 QY 2567 TTTGTGTCATCAGACAGACAGGATTTAAATATTTAACTTATTTTAACTTAACTTAACT 2626  
 DB 4131 TTTGTGTCATCAGACAGACAGGATTTAAATATTTAACTTATTTTAACTTAACTTAACT 4190  
 QY 2627 AGAAGGAATCCATTCCTAGCTTTCTGTTGTTGCTCTAAATTTTGGTGGTGGTGGG 2686  
 DB 4191 AGAAGGAATCCATTCCTAGCTTTCTGTTGTTGCTCTAAATTTTGGTGGTGGTGGG 4250  
 QY 2687 ATCCCAACAATCAGTCCCTGAGATAGCTGCTATTTGGGCTGATTCATTGCGCAGATCT 2746  
 DB 4251 ATCCCAACAATCAGTCCCTGAGATAGCTGCTATTTGGGCTGATTCATTGCGCAGATCT 4310  
 QY 2747 TCTTCTCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2806  
 DB 4311 TCTTCTCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4370  
 QY 2807 TCCCAATATGATTTCCAAATGCTTTACCAAGTTTAGGGTGTTCGAAGAGGTAGAGG 2866  
 DB 4371 TCCCAATATGATTTCCAAATGCTTTACCAAGTTTAGGGTGTTCGAAGAGGTAGAGG 4430  
 QY 2867 GTGGGGCTTCCAGCTCTCAACGGCTTCCCTTAACCAACCCCTTCTCTTTGGCCAGCTGT 2926  
 DB 4431 GTGGGGCTTCCAGCTCTCAACGGCTTCCCTTAACCAACCCCTTCTCTTTGGCCAGCTGT 4490  
 QY 2927 TCCCCCACTTCCACTCCCTCTACTCTCTTAGGACTGGGCTGATGAAGGCACTGCCCA 2986  
 DB 4491 TCCCCCACTTCCACTCCCTCTACTCTCTTAGGACTGGGCTGATGAAGGCACTGCCCA 4550  
 QY 2987 AAATTTCCCTTACCCCACTTCCCTTACCCCACTTCCCTTACCCCACTTCCCTTACCCCA 3046  
 DB 4551 AAATTTCCCTTACCCCACTTCCCTTACCCCACTTCCCTTACCCCACTTCCCTTACCCCA 4610  
 QY 3047 TGTGTTGAGTCTGTCAGGAGCAGACAGCAAAAGTGGGTTTCCCAAGCCCTTTGTCATC 3106  
 DB 4611 TGTGTTGAGTCTGTCAGGAGCAGACAGCAAAAGTGGGTTTCCCAAGCCCTTTGTCATC 4670  
 QY 3107 TCAGCCCGCAGCTATATCTGCTTGGGAACTCTACACAGAACTCAGAGCAGCCCC 3166  
 DB 4671 TCAGCCCGCAGCTATATCTGCTTGGGAACTCTACACAGAACTCAGAGCAGCCCC 4730  
 QY 3167 TGCCTGAGCTAAGGAGGCTCTTATCTCTCAGGGGGGTTTAAAGTCCGTTTGAATAATG 3226  
 DB 4731 TGCCTGAGCTAAGGAGGCTCTTATCTCTCAGGGGGGTTTAAAGTCCGTTTGAATAATG 4790  
 QY 3227 TCGTCTTATTTATTTAGCGGGTGAATATTTTATCTGTAAGTGAAGCAATCAGAGTATAA 3286  
 DB 4791 TCGTCTTATTTATTTAGCGGGTGAATATTTTATCTGTAAGTGAAGCAATCAGAGTATAA 4850  
 QY 3287 TGTGTTGCTGACAAAATTAAGGCTTCTTATATGTTTAAAAA 3330  
 DB 4851 TGTGTTGCTGACAAAATTAAGGCTTCTTATATGTTTAAAAA 4894

RESULT 23

AAD05230  
 ID AAD05230 standard; cdna; 3878 BP.

XX AAD05230;

XX AC

XX DT

XX DE

18-JUL-2001 (first entry)

Human secreted protein-encoding gene 11 cdna clone HWBARI4, SEQ ID NO:21.



XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 KW cell culture; chemotaxis; vulnery; binding partner identification;  
 KW gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 152..1267

XX FT /\*tag= a

XX FT /product= "Human secreted protein"

XX FT /transl\_except= (pos:209..211, aa:Xaa)

XX FT /note= "Xaa corresponds to any of the naturally occurring

XX FT l-amino acids"

XX FT sig\_peptide 152..295

XX FT /\*tag= b

XX FT mat\_peptide 296..1264

XX FT /\*tag= c

XX FT /product= "Mature human secreted protein"

XX WO200134629-A1.

XX 17-MAY-2001.

XX 08-NOV-2000; 2000WO-US30654.

XX 12-NOV-1999; 98US-0164835.

XX 27-JUL-2000; 2000US-0221142.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis GA, Wei P, Fiscella M, Baker KP;

XX WPI: 2001-308779/32.

XX P-PSDB; AAE01362.

XX New nucleic acid encoding one of 21 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions,  
 PT such as autoimmune disease and cancer, and used as a food additive or  
 PT preservative

XX Claim 1; Page 388-389; 490pp; English.

XX AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted  
 CC protein genes, and AAE01352-AAE01413 represent the proteins they encode.  
 CC AAE01415-AAE01433 represent human secreted protein fragments or variants.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 21 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infections. The  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or

CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
 CC The present sequence represents a human secreted protein-encoding cDNA of  
 CC the invention.

XX SQ Sequence 3878 BP; 709 A; 1164 C; 1044 G; 951 T; 10 other;

Query Match 62.78; Score 2136.8; DB 22; Length 3878;

Best Local Similarity 80.0%; Pred. No. 0;

Matches 2885; Conservative 5; Mismatches 4; Indels 710; Gaps 4;

Qy 455 GGCATTGCTCCAGTGGCTGGCTGGTGTGTCCTCCGCTCTAGCTCAGCAGTGCAC 514

Db 245 GGCATTGCTCCAGTGGCTGGCTGGTGTGTCCTCCGCTCTAGCTCAGCAGTGCAC 304

Qy 515 TGGCGTGAGCGCTATGATGCCGCCGCCGCCCTTCATCTGGGCACTGCTTGGGCATCCTG 574

Db 305 TGGCGWAGAGCTATGATGCCGCCGCCGCCCTTCATCTGGGCACTGCTTGGGCATCCTG 364

Qy 575 CTGAGCCTCTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGCTTGGCCGATCCC 634

Db 365 CTGAGCCTCTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGCTTGGCCGATCCC 424

Qy 635 AGGCCCTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGCTGACTTCTTGGCCAG 694

Db 425 AGGCCCTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGCTGACTTCTTGGCCAG 484

Qy 695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTGACCTCTTCGGGACCCGGACCACTGT 754

Db 485 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTGACCTCTTCGGGACCCGGACCACTGT 544

Qy 755 CGCAGGCGCTACTCTGCTATGCTTTCATGATCAGTCTTGGGGCTGCTGGGCTACCTC 814

Db 545 CGCAGGCGCTACTCTGCTATGCTTTCATGATCAGTCTTGGGGCTGCTGGGCTACCTC 604

Qy 815 CTGCTTGCATTTGCTGCTGGACCACTGCTGCCCCCTTACCTGGGCACTGGGCACTGCTG 874

Db 605 CTGCTTGCATTTGCTGCTGGACCACTGCTGCCCCCTTACCTGGGCACTGGGCACTGCTG 664

Qy 875 TGCTCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 934

Db 665 TGCTCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724

Qy 935 GCTGAGGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 994

Db 725 GCTGAGGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784

Qy 995 TCGCCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1054

Db 785 TCGCCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844

Qy 1055 CCGCGGCTGCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1114

Db 845 CCGCGGCTGCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 904

Qy 1115 GAGCTGTGAGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1174

Db 905 GAGCTGTGAGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 964

Qy 1175 GAGGGGCTGTACAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1234

Db 965 GAGGGGCTGTACAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1024

Qy 1235 GATGA----- 1239

Db 1025 GATGAAGGTAAAGGCTTGGACAG 1084

Qy 1240 ----- 1239

Db 1085 AACTCGGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1144



QY	1240	-----	1239	:	2225	GGTTMMGGGGGGGGG	2284		
Db	1145	AAAGTGGGGATGGACCCCATCTGCAVACACGGCTTCTCATGAGGTGTGGAACATCTCTGCT	1204		QY	1789	CCAGGTGGCCCCCATCCCTGTTTATGGGCTCCATFTGCCAGCTCAGCAGCTCTGTCTACCTGC	1848	
QY	1240	-----	1239		Db	2285	CCAGGTGGCCCCCATCCCTGTTTATGGGCTCCATFTGCCAGCTCAGCAGCTCTGTCTACCTGC	2344	
Db	1205	TGCGGTFTTCAGGAAGGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCTGTGGCCCGAGTT	1264		QY	1849	CTATATGGTGTCTGCGCGCAGGCCTGGGTCTGGTGCCTATTTACTTTGTCTACACAGTAGT	1908	
QY	1240	-----	1239		Db	2345	CTATATGGTGTCTGCGCGCAGGCCTGGGTCTGGTGCCTATTTACTTTGTCTACACAGTAGT	2404	
Db	1265	TCACAGAGGAAGGGCGAGCTTATTTCAAAGTCTAGAGGAGTGGAGGAGTTAAAGCTGG	1324		QY	1909	ATTTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACCTTCCAGCACATTTGGGGTG	1968	
QY	1240	-----	1239		Db	2405	ATTTGNCAGAAGAGCGACTTGG-CAATACTCAGCGTAG-AAACTTCCAGCMCAATTTGGGTG	2462	
Db	1325	ATTTTCAGATCTGCCTGTTTCCAGCCGAGTGTGCCCTCTGCTCCCCCAACAGACTTTTCCAA	1384		QY	1969	GAGGGCTTGCCTCACTGGGTCCCAGGCTCCCGCTCTCTGTGTAGCCCATGGGCTGCCGG	2028	
QY	1240	-----	1239		Db	2463	RAGGGCTGCCTCACTGGGTCCCAGGCTCCCGCTCTCTGTGTAGCCCATGGGCTGCCGG	2522	
Db	1385	ATAATCTCACAGCGCCTTCCAGCTCAGCGTCTAGAGCGTCTTAGAAGCGTCTTGAAGCCTATGGCCA	1444		QY	2029	CTGCGCGCGAGTTTCTGTTGTCGCCAAAGTAATGTGGCTCTCTGTCTGTGCCACCTGTGTCTG	2088	
QY	1240	-----	1239		Db	2523	CTGCGCGCGAGTTTCTGTTGTCGCCAAAGTAATGTGGCTCTCTGTCTGTGCCACCTGTGTCTG	2582	
Db	1445	GCTGTCTTTGTTCCTCTCACCCGCCCTGTCTCACAGCTGAGACTCCCAGGAACTT	1504		QY	2089	CTGAGGTGCGTAGCTGCACAGCTGGGGCTGGGGCTCCCTCTCTCTCTCCCGACGTCTC	2148	
QY	1240	-----	1239		Db	2583	CTGAGGTGCGTAGCTGCACAGCTGGGGCTGGGGCTCCCTCTCTCTCTCTCCCGACGTCTC	2642	
Db	1505	CAGACTACCTTCTCTGCTTTCAGCAAGGGCGTTGCCACATTTCTCTGAGGGTCAGTGG	1564		QY	2149	TAGGCTTGCCTTACTTGGAGGCTTCCAAAGGGGTTTTCAGTCTGAGCTTATACAGGGAGGC	2208	
QY	1240	-----	1239		Db	2643	TAGGCTTGCCTTACTTGGAGGCTTCCAAAGGGGTTTTCAGTCTGAGCTTATACAGGGAGGC	2702	
Db	1565	AAGAACTTAGACTCCCATTTGCTAGAGGTAGAAAGGGGAGGGTGTGGGGAGCAGGGTG	1624		QY	2209	CAGAAAGGCTCCATGTGACTTGAATGCGGGGACTCTGCAGTGTGATTTACCCAGGCTCAGGG	2268	
QY	1240	-----	1239		Db	2703	CAGAAAGGCTCCATGTGCACTGGAATGCGGGGACTCTGCAGGTGGATTTACCCAGGCTCAGGG	2762	
Db	1625	GTCCACACGAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGGCTCTCATCTCCCTGAGAC	1684		QY	2269	TTACAGCTAGCCCTCTAGTTTGAGACACACCTAGAGAAGGTTTGGGAGCTGAATAAA	2328	
QY	1240	-----	1248	-----AGCGTTTC	Db	2763	TTACAGCTAGCCCTCTAGTTTGAGACACACCTAGAGAAGGTTTGGGAGCTGAATAAA	2822	
Db	1685	TGCTCCGACCTTCCCTCCAGGCTCTGTCTGATGGCCCTCTCCCTCTCGAGGGCGTTTCG	1744		QY	2329	CTCAGTCACTTGGTTTCCCATCTCTAAGCCCTTTAAGCTCAGCTTCGTTTAAATGTAGCT	2388	
QY	1249	GATGGCAGCCTGGGCTGTTCTTGCACTGGCGCATCTCCCTGGTCTTCTCTCTGCTCAT	1308		Db	2823	CTCAGTCACTTGGTTTCCCATCTCTAAGCCCTTTAAGCTCAGCTTCGTTTAAATGTAGCT	2882	
Db	1745	GATGGCAACCTTGGGGCTGTTCTTGCACTGGCGCATCTCCCTGGTCTTCTCTGCTCAT	1804		QY	2389	CTTGATGGGAGTTTCTAGGATGAACACATCTCCATCGGATTTGAACATATG--ACTTA	2446	
QY	1309	GGACGGCTGGTGACGAGATTTCGGCACTTCGACAGTCTATTGCGCAGTGTGGCAGCTTT	1368		Db	2883	CTTGATGGGAGTTTCTAGGATGAACACATCTCCATCGGATTTGAACATATGAAGTTA	2942	
Db	1805	GGACCGGCTGGTGACAGCATTCGGCACTCGACAGTCTATTGCGCAGTGTGGCAGCTTT	1864		QY	2447	TTTGTAGGGAGAGTCTCTGAGGGCACACACAGAACCAGGTCCCTCAGCCACACAGC	2506	
QY	1369	CCCTGTGGCTGCCGTGCCACATGCTGTCCACAGTGTGGCGTGGTGACAGTTTCAGC	1428		Db	2943	TTTGTAGGGAGAGTCTCTGAGGGCACACACAGAACCAGGTCCCTCAGCCACAGC	3002	
Db	1865	CCCTGTGGCTGCCGTGCCACATGCTGTCCACAGTGTGGCGTGGTGACAGCTTCAGC	1924		QY	2507	ACTGTCTTTTGTGATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTGTGGTCC	2566	
QY	1429	CGCCCTCACCGGTTACCTTCTCAGCCCTGCAGATCTGCCCTTACACACTGGCCCTCCT	1488		Db	3003	ACTGTCTTTTGTGATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTGTGGTCC	3062	
Db	1925	CGCCCTCACCGGTTACCTTCTCAGCCCTGCAGATCTGCCCTTACACACTGGCCCTCCT	1984		QY	2567	TTCTGTGCGCATCACAGAGACAGGCAATTTAAATATTTAACTTATTTTAAACAAAGT	2626	
QY	1489	CTACACCGGGAGAGCAGGTGTTCTGCCCAATACCGAGGGGACACTGGAGGTGCTAG	1548		Db	3063	TTCTGTGCGCATCACAGAGACAGGCAATTTAAATATTTAACTTATTTTAAACAAAGT	3122	
Db	1985	CTACACCGGGAGAGCAGGTGTTCTGCCCAATACCGAGGGGACACTGGAGGTGCTAG	2044		QY	2627	AGAAAGGAATCCATTTGCTAGCTTTTCTGTGTGTGTCTAATATTTGGTAGGGTGGGG	2686	
QY	1549	CAGTGAGGACGCCCTGATGACAGTCTCTGCCAGGCCCTAAGCCTGAGAGTCCCTTCCC	1608		Db	3123	AGAAAGGAATCCATTTGCTAGCTTTTCTGTGTGTGTCTAATATTTGGTAGGGTGGGG	3182	
Db	2045	CAGTGAGGACGCCGTATGACCAGCTTCTGCCAGGCCCTAAGCCTGAGCTCCCTTCCC	2104		QY	2687	ATCCCCCAACATCAGGTCCCTGAGATAGCTGGTCTATTT		



1837	QY	GTCTCTCACTGCCTATATGGTGTCTGCGCGCAGGCGCTGGGTCTGGTCTGCTTACTTTGC	1896
1838	QY		
1839	QY		
1840	QY		
661	Db	GTCGTGCTACTGCCATATATGGTGTCTGCGCAGGCGCTGGTCTGGTCTGCGCAATTTACTTTGC	720
1897	QY	TACACAGGTAGTATTTGACACAGAGCGACTTGGCCAAATACTACAGCGTAGAAAACTTCAG	1956
721	Db	TACACAGGTAGTATTTGACACAGAGCGACTTGGCCAAATACTACAGCGTAGAAAACTTCAG	780
1957	QY	CACATTGGGGTGGAGGGCTGCCTCACATGGGTGCCAGCTCCCGCGTCTCTGTTAGCCCAT	2016
781	Db	CACATTGGGGTGGAGGGCTGCCTCACATGGGTGCCAGCTCCCGCGTCTCTGTTAGCCCAT	840
2017	QY	GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC	2076
841	Db	GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC	900
2077	QY	CACCTGTGCTGCTCAGGTGCGTAGCTGCACAGCTGGGGCTGGGGGCTGCCCTCTCCCTCT	2136
901	Db	CACCTGTGCTGCTCAGGTGCGTAGCTGCACAGCTGGGGCTGGGGGCTGCCCTCTCCCTCT	960
2137	QY	CTCCCCAGTCTCTAGGGTGCTGACTGGAGGCGTTCCAAGGGGTTTCAGTCTGGACTT	2196
961	Db	CTCCCCAGTCTCTAGGGTGCTGACTGGAGGCGTTCCAAGGGGTTTCAGTCTGGACTT	1020
2197	QY	ATACAGGAGGCGCAGAAGGGCTCCATGCACCTGGAATGCGGGGACTCTGCAGTGGATTAC	2256
1021	Db	ATACAGGAGGCGCAGAAGGGCTCCATGCACCTGGAATGCGGGGACTCTGCAGTGGATTAC	1080
2257	QY	CGAGCTCAGGGTTTACAGCTAGCCTCCTAGTTTGACACACCTAGAGAAAGGTTTGG	2316
1081	Db	CGAGCTCAGGGTTTACAGCTAGCCTCCTAGTTTGACACACCTAGAGAAAGGTTTGG	1140
2317	QY	GAGCTGAATAAACTCAGTCACCTGGTTTCCCCTCTCTAAGCCCCCTAACTCAGAGTTGG	2376
1141	Db	GAGCTGAATAAACTCAGTCACCTGGTTTCCCCTCTCTAAGCCCCCTAACTCAGAGTTGG	1200
2377	QY	TTTAAATGAGTCTTGTCATGGAGGTTTCTAGGATGAACACATCTCTCCATGGGATTTGAAC	2436
1201	Db	TTTAAATGAGTCTTGTCATGGAGGTTTCTAGGATGAACACATCTCTCCATGGGATTTGAAC	1260
2437	QY	ATATG - ACTTATTTGTAGGGGAAGATGCTCTGAGGGCAACACACAAGAACCAAGTCCCC	2494
1261	Db	ATATGAAGTATTTGTAGGGGAAGATGCTCTGAGGGCAACACACAAGAACCAAGTCCCC	1320
2495	QY	TCAGCCCACAGCACTGTCTTTTGTGTATCCACCCGCCCTCTTACCTTTTATCAGGATGTG	2554
1321	Db	TCAGCCCACAGCACTGTCTTTTGTGTATCCACCCGCCCTCTTACCTTTTATCAGGATGTG	1380
2555	QY	GCCTGTTGTCCTCTGTTGCCATCACACAGACACAGGCGATTAAATATTTAACTTATTT	2614
1381	Db	GCCTGTTGTCCTCTGTTGCCATCACACAGACACAGGCGATTAAATATTTAACTTATTT	1440
2615	QY	ATTTAAACAAGTAGAAGGGAATCCATGTCTAGCTTTCTGTGTGGTGTCTAATAATTGG	2674
1441	Db	ATTTAAACAAGTAGAAGGGAATCCATGTCTAGCTTTCTGTGTGGTGTCTAATAATTGG	1500
2675	QY	GTAGGTGGGGGATCCCCAACAAATCAGGTCCCCCTGAGATAGCTGTGTCATTTGGGCTGATCA	2734
1501	Db	GTAGGTGGGGGATCCCCAACAAATCAGGTCCCCCTGAGATAGCTGTGTCATTTGGGCTGATCA	1560
2735	QY	TTGCCAGAATCTTCTTCTCTCTGGGTCTGGCCCCCCCCAAATGCCTAACCCAGGACCTGG	2794
1561	Db	TTGCCAGAATCTTCTTCTCTCTGGGTCTGGCCCCCCCCAAATGCCTAACCCAGGACCTGG	1620
2795	QY	AAATTCTACTCATCCCAAAATGATAATTCCAATATGCTGTTACCCAAGGTTAGGCTGTGAA	2854
1621	Db	AAATTCTACTCATCCCAAAATGATAATTCCAATATGCTGTTACCCAAGGTTAGGCTGTGAA	1680
2855	QY	GGAAGGTAGAGGTGGGGCTTTCAGGTCTCAACGGCTTCCCTTAACACCCCTCTTCTCTTG	2914
1681	Db	GGAAGGTAGAGGTGGGGCTTTCAGGTCTCAACGGCTTCCCTTAACACCCCTCTTCTCTTG	1740

Qy	2915	GCCACGCGTGGTTCCCGCCCACTTCCACTCCCTCTCTACTCTCTCTAGACTGGCGCTGATGA	2974
Db	1741	GCCACGCGTGGTTCCCGCCCACTTCCCACTCCCTCTCTACTCTCTCTAGACTGGCGCTGATGA	1800
Qy	2975	AGCACTGCCCAAAATTTCCCTTACCCCAAACTTTCCCTTACCCCAAACTTTCCCGCCACCA	3034
Db	1801	AGCACTGCCCAAAATTTCCCTTACCCCAAACTTTCCCTTACCCCAAACTTTCCCGCCACCA	1860
Qy	3035	GCTCCCAACCCCTGTTGGAGCTACTGCAGGACCAAGAAGTGCAGTTCCTCCCAAG	3094
Db	1861	GCTCCCAACCCCTGTTGGAGCTACTGCAGGACCAAGAAGTGCAGTTCCTCCCAAG	1920
Qy	3095	CCTTTGTCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAACTC	3154
Db	1921	CCTTTGTCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAACTC	1980
Qy	3155	AGGAGCACCCCTGCCTGAGCTAAGGAGGCTCTTATCTCTCAGGGGGGTTTAAGTCCGC	3214
Db	1981	AGGAGCACCCCTGCCTGAGCTAAGGAGGCTCTTATCTCTCAGGGGGGTTTAAGTCCGC	2040
Qy	3215	TTTGCAATAATGCGTCTTATTTATTTAGCGGGTGAATATTTTATATCTGTAAGTGAGCA	3274
Db	2041	TTTGCAATAATGCGTCTTATTTATTTAGCGGGTGAATATTTTATATCTGTAAGTGAGCA	2100
Qy	3275	ATCAGAGTAAATGTTTTATGTCGACAAAATTAAGGCTTCTTATATGTTTA	3326
Db	2101	ATCAGAGTAAATGTTTTATGTCGACAAAATTAAGGCTTCTTATATGTTTA	2152
RESULT	25		
AAV71180			
ID	AAV71180	standard; cDNA; 2143 BP.	
XX	AC	AAV71180;	
XX	AC		
DT	12-FEB-1999	(first entry)	
XX	XX		
DE	XX	Clone 1711346IH, the PS108 gene contig full length sequence.	
XX	XX		
KW	KW	PS108 gene; prostate disease; benign prostatic hyperplasia; BPH;	
KW	KW	prostatitis; prostatic intraepithelial neoplasia; PIN; cancer;	
KW	KW	drug screening; gene therapy; ss.	
XX	OS	Homo sapiens.	
XX	OS		
PN	PN	W09850567-A1.	
XX	XX		
PD	PD	12-NOV-1998.	
XX	XX		
PF	PF	01-MAY-1998; 98WO-US08930.	
XX	XX		
PR	PR	02-MAY-1997; 97US-0850713.	
XX	XX		
PA	PA	(ABBO ) ABBOTT LAB.	
XX	XX		
PI	PI	Billings-Medel PA, Cohen M, Colpitts TL, Friedman PN;	
PI	PI	Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;	
PI	PI	Roberts-Rapp L, Russell JC, Stroupe SD;	
XX	XX		
DR	DR	WPI; 1999-034731/03.	
XX	XX		
PT	PT	New isolated prostate-specific polynucleotides - used to develop	
PT	PT	products for the diagnosis and treatment of prostate diseases, e.g.	
PT	PT	benign hyperplasia, prostatic or prostate cancer	
XX	XX		
PS	PS	Claim 1; Fig 1A-E; 122pp; English.	
XX	XX		
CC	CC	The present sequence represents the full length contig of the PS108	
CC	CC	gene, as represented by clone 1711346IH. This clone is the contig of	
CC	CC	overlapping clones AAV71166-79. The clone sequences are PS108	
CC	CC	gene-specific. They are used in the method of the invention. The	
CC	CC	specification describes a method for detecting the presence of a	
CC	CC	target PS108 polynucleotide in a test sample. The method comprises	



QY 3163 CCCTGCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAGTCCGCTTTGCAAT 3222  
 DB 1980 CCCCTGCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAGTCCGCTTTGCAAT 2039  
 QY 3223 AATGCTGCTTATTTATTTAGCGGGTGAATATTTTATCTGTAAGTGAAGCAATCAGACT 3282  
 DB 2040 AATGCTGCTTATTTATTTAGCGGGTGAATATTTTATCTGTAAGTGAAGCAATCAGACT 2099  
 QY 3283 ATAATCTTTATGTCGACAAATTTAAAGGCTTTCTTATATGTTTA 3326  
 DB 2100 ATAATGTTTATGTCGACAAATTTAAAGGCTTTCTTATATGTTTA 2143

RESULT 26

AZ45677  
 ID AZ45677 standard; cDNA; 2462 BP.

XX AC AZ45677;

XX 06-APR-2000 (first entry)

DE cDNA sequence of a novel prostate cancer-associated gene.

XX Prostate cancer-associated gene; Incyte clone 1864683; bone cancer;  
 KW cell proliferation; cancer; adrenal gland cancer; bladder cancer;  
 KW prostate cancer; ss.

OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT CDS 376..1071  
 FT /\*tag= a

FT /\*product= "prostate cancer-associated protein"

FT sig\_peptide 376..516

FT /\*tag= b

XX W09967384-A2.

XX 29-DEC-1999.

XX 15-JUN-1999; 99WO-US13524.

XX 22-JUN-1998; 98US-0102615.

XX (INCY-) INCYTE PHARM INC.

XX Walker MG, Volkmath W, Klingler TM, Sprinzak EA;

XX WPI; 2000-126631/11.

XX P-PSDB; AAY54369.

XX Identifying biomolecules for the diagnosis and treatment of diseases  
 associated with cell-proliferation -

XX Claim 8; Page 47-48; 52pp; English.

XX The present sequence is derived from a prostate cancer-associated  
 CC gene, and is represented by Incyte clone number 1864683. The sequence  
 CC is used in the method of the invention. The specification  
 CC describes a method for identifying biomolecules for the diagnosis or  
 CC treatment of diseases associated with cell proliferation. The method  
 CC comprises examining polynucleotides, consisting of prostate cancer-  
 CC specific genes, and genes of unknown function, expressed in cDNA  
 CC libraries. The patterns of both gene sets are compared to identify  
 CC genes of unknown function with similar expression patterns to the  
 CC prostate cancer-specific genes. The biomolecules identified by the  
 CC method form pharmaceutical compositions useful for the diagnosis and  
 CC treatment of diseases associated with cell proliferation. Such diseases  
 CC include cancer of the adrenal gland, bladder and bone, but especially  
 CC prostate cancer. The method may also be applied using other disease-  
 CC specific genes. The prostate cancer-specific genes facilitate the  
 CC diagnosis and treatment of cell-proliferation disorders.

XX SQ Sequence 2462 BP; 483 A; 726 C; 632 G; 621 T; 0 other;  
 Query Match 60.6%; Score 2065.4; DB 21; Length 2462;  
 Best Local Similarity 99.6%; Pred. No. 7.6e-314;  
 Matches 2092; Conservative 0; Mismatches 6; Indels 3; Gaps 2;  
 QY 1237 TGAAGCGCTTCGGATGGGAGCGCTGGGCTGTCTCTCAGTGGCGCATCTCCCTGGTCTT 1296  
 DB 363 TGCAGCGCTTCGGATGGGAGCGCTGGGCTGTCTCTCAGTGGCGCATCTCCCTGGTCTT 422  
 QY 1297 CTCTCTGCTCATGAGACCGGCTGGTGCAGCGATTCGGCACATCGAGCAGTCTATTTGGCCAG 1356  
 DB 423 CTCTCTGCTCATGAGACCGGCTGGTGCAGCGATTCGGCACATCGAGCAGTCTATTTGGCCAG 482  
 QY 1357 TGTGGCAGCTTTCCTCTGTGGTGGCGGTGCCATGCTGCCACATGTGGCGGTGGT 1416  
 DB 483 TGTGGCAGCTTTCCTCTGTGGTGGCGGTGCCATGCTGCCACATGTGGCGGTGGT 542  
 QY 1417 GACAGCTTCAGCGCGCTCACCGGCTTCACCTTCTCAGCCCTGCAGATCTCCCTACAC 1476  
 DB 543 GACAGCTTCAGCGCGCTCACCGGCTTCACCTTCTCAGCCCTGCAGATCTCCCTACAC 602  
 QY 1477 ACTGCGCTCCCTCTTACCACCGGGAGAGCAGGTGTTCTGCCCAATACCAGAGGGACAC 1536  
 DB 603 ACTGCGCTCCCTCTTACCACCGGGAGAGCAGGTGTTCTGCCCAATACCAGAGGGACAC 662  
 QY 1537 TGGAGGTGCTAGCAGTGGAGCAGCTGATGACACAGCTTCTGCCAGCGCTTACGCTGG 1596  
 DB 663 TGGAGGTGCTAGCAGTGGAGCAGCTGATGACACAGCTTCTGCCAGCGCTTACGCTGG 722  
 QY 1597 AGCTCCCTTCCCTAATGGACAGCTGGGTGCTGGAGGAGTGGCTTCCACCTCCACACC 1656  
 DB 723 AGCTCCCTTCCCTAATGGACAGCTGGGTGCTGGAGGAGTGGCTTCCACCTCCACACC 782  
 QY 1657 CGCGCTCTCGGGGCTCTGCGCTGTGATGCTCTCCGTACGTGTGTGGTGGTGGTGGTGG 1716  
 DB 783 CGCGCTCTCGGGGCTCTGCGCTGTGATGCTCTCCGTACGTGTGTGGTGGTGGTGGTGG 842  
 QY 1717 CGAGGCCAGGTGGTTCGGGCGGGGCGGATCTGCTGGACCTCGGCATCTCGGATAGTGC 1776  
 DB 843 CGAGGCCAGGTGGTTCGGGCGGGGCGGATCTGCTGGACCTCGGCATCTCGGATAGTGC 902  
 QY 1777 CTCTCTGCTGCCAGGTGGCGCCCATCCCTGTTTATGGCTCCATGTGCCAGCTCAGCCA 1836  
 DB 903 CTCTCTGCTGCCAGGTGGCGCCCATCCCTGTTTATGGCTCCATGTGCCAGCTCAGCCA 962  
 QY 1837 GTCTGTCACTGCCTATATGTTGCTGTGCCGAGCGCTGGGTCTGTGGTGGTGGTGGTGG 1896  
 DB 963 GTCTGTCACTGCCTATATGTTGCTGTGCCGAGCGCTGGGTCTGTGGTGGTGGTGGTGG 1022  
 QY 1897 TACACAGTGTATTTGACAGAGCGACTTGGCCAAATACTCAGCGTAGAAGCTTCCAG 1956  
 DB 1023 TACACAGTGTATTTGACAGAGCGACTTGGCCAAATACTCAGCGTAGAAGCTTCCAG 1082  
 QY 1957 CACATTGGGGTGGAGGGCTGCTCACTGGGTGCCAGCTTCCCGCTCCCTGTAGCCCAT 2016  
 DB 1083 CACATTGGGGTGGAGGGCTGCTCACTGGGTGCCAGCTTCCCGCTCCCTGTAGCCCAT 1142  
 QY 2017 GGGGTGCGGGCTGGCGCCAGTGTCTGTCTGCCAAAGTAATGTGGCTCTCTGTGCTGTC 2076  
 DB 1143 GGGGTGCGGGCTGGCGCCAGTGTCTGTCTGCCAAAGTAATGTGGCTCTCTGTGCTGTC 1202  
 QY 2077 CACCCTGTCTCTGAGGTGCTGCTGACAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 2136  
 DB 1203 CACCCTGTCTCTGAGGTGCTGCTGACAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 1261  
 QY 2137 CTCGCCAGTCTCTAGGGCTGCTGACTGGAGCGCTTCCAAAGGGGCTTTCAGTCTGACTT 2196  
 DB 1262 CTCGCCAGTCTCTAGGGCTGCTGACTGGAGCGCTTCCAAAGGGGCTTTCAGTCTGACTT 1321  
 QY 2197 ATACAGGGAGCGCAGAGGGCTTCCATGCACTGGAATGGGGGACTCTCGAGGTGGATTCAC 2256

Db 1322 ATACAGGGAGCCAGAGGGCTCCATGCACTGGAATGCGGGACTCTCGAGGTGATAC 1381  
QY 2257 CCAGGCTCAGGTTTAAACAGCTAGCCTCTAGTGTGAGACACACCTAGAGAAGGGTTTGTG 2316  
Db 1382 CCAGGCTCAGGTTTAAACAGCTAGCCTCTAGTGTGAGACACACCTAGAGAAGGGTTTGTG 1441  
QY 2317 GAGCTGAATAAAGTCACTGAGTCACTGCTGTTCCCATCTCTAAGCCCTTAACCTGCAGCTTCG 2376  
Db 1442 GAGCTGAATAAAGTCACTGAGTCACTGCTGTTCCCATCTCTAAGCCCTTAACCTGCAGCTTCG 1501  
QY 2377 TTTAATGTAGCTCTTGCATGGGAGTCTTAGGATGAACACATCTCTCCATGGATTTTGAAC 2436  
Db 1502 TTTAATGTAGCTCTTGCATGGGAGTCTTAGGATGAACACATCTCTCCATGGATTTTGAAC 1561  
QY 2437 ATATG--ACTTATTTGTAGGGAAGAGTCTGTAGGGAACACACACAAAGACAGTCCCG 2494  
Db 1562 ATATGAAGTTTATTTGTAGGGAAGAGTCTGTAGGGAACACACACAAAGACAGTCCCG 1621  
QY 2495 TCAGCCACAGCACTGTCTTTTGTGTCATGATCCACCCCTCTTACCTTTTATCAGATGTG 2554  
Db 1622 TCAGCCACAGCACTGTCTTTTGTGTCATGATCCACCCCTCTTACCTTTTATCAGATGTG 1681  
QY 2555 GCCTGTGTCTCTCTGTTGCCATCACAGACACAGGCAATTTAAATATTTAACTTATTT 2614  
Db 1682 GCCTGTGTCTCTCTGTTGCCATCACAGACACAGGCAATTTAAATATTTAACTTATTT 1741  
QY 2615 ATTTAAAGTGTAGGGAAGTCTTCTGTAGCTTTTCTGTGTTGGTCTCTAATTTTGG 2674  
Db 1742 ATTTAAAGTGTAGGGAAGTCTTCTGTAGCTTTTCTGTGTTGGTCTCTAATTTTGG 1801  
QY 2675 GTAGGTTGGGGATCCCAACATCAGTCCCTGAGATAGCTGTCTATTTGGCTGATCA 2734  
Db 1802 GTAGGTTGGGGATCCCAACATCAGTCCCTGAGATAGCTGTCTATTTGGCTGATCA 1861  
QY 2735 TTGCAGAACT 2794  
Db 1862 TTGCAGAACT 1921  
QY 2795 AAATCTACTATCCCAATGATAATCCAAATGCTGTACCCAAAGTGTAGGTTGTTGAA 2854  
Db 1922 AAATCTACTATCCCAATGATAATCCAAATGCTGTACCCAAAGTGTAGGTTGTTGAA 1981  
QY 2855 GAAGGTAGAGGTGGGCTCAGCTCTCAAGGCTTCCCTTAACCAACCCCTCTCTCTCT 2914  
Db 1982 GGAAGTGTAGGTTGGGCTCAGCTCTCAAGGCTTCCCTTAACCAACCCCTCTCTCTCT 2041  
QY 2915 GCCAGCTGTGTCT 2974  
Db 2042 GCCAGCTGTGTCT 2101  
QY 2975 AGGCACTGCCAAATTTCCCTACCCCAACTTTCCCTACCCCAACTTTCCCACTCA 3034  
Db 2102 AGGCACTGCCAAATTTCCCTACCCCAACTTTCCCTACCCCAACTTTCCCACTCA 2161  
QY 3035 GCTCCACAACTGTGTGGAGTCTCAGACACAGAGCAAGTGTGGTTTCCCAAG 3094  
Db 2162 GCTCCACAACTGTGTGGAGTCTCAGACACAGAGCAAGTGTGGTTTCCCAAG 2221  
QY 3095 CCTTTGTCCATCTCAGCCCCAGAGTATATCTGTCTTGGGAATCTCACACAAACTC 3154  
Db 2222 CCTTTGTCCATCTCAGCCCCAGAGTATATCTGTCTTGGGAATCTCACACAAACTC 2281  
QY 3155 AGGACACCCCTGTCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAAGTCCG 3214  
Db 2282 AGGACACCCCTGTCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAAGTCCG 2341  
QY 3215 TTGCAATTAAGTGTCTTATTTATTTAGCGGGGTGAATATTTATCTGTAAGTGAGCA 3274  
Db 2342 TTGCAATTAAGTGTCTTATTTATTTAGCGGGGTGAATATTTATCTGTAAGTGAGCA 2401  
QY 3275 ATCAGATATTAAGTGTGTGACAAATTAAGGCTTCTTATATCTTTTAAAAA 3334  
Db 2402 ATCAGATATTAAGTGTGTGACAAATTAAGGCTTCTTATATATGTTAAAAA 2461

QY 3335 A 3335  
Db 2462 A 2462

## RESULT 27

AAC84928  
ID AAC64928 standard; DNA; 2133 BP.

XX AAC64928;

XX 07-FEB-2001 (first entry)

XX Human prostate-related PS108 partial coding sequence SEQ ID NO: 16.

XX Human; prostate cancer; PS108; antibody; tumour; metastasis; ds.

OS Homo sapiens.

XX US6130043-A.

XX 10-OCT-2000.

XX 01-MAY-1998; 98US-0071710.

XX 02-MAY-1997; 97US-0850713.

XX (ABBO ) ABBOTT LAB.

XX Cohen M, Colpitts TL, Friedman PN, Gordon J, Granados EN;  
PI Billings-Medel PA, Klass MR, Roberts-Rapp L, Stroupe SD, Yu H;  
PI Kratochvil JD, Russell JC, Hodges SC;

XX WPI; 2000-655655/63.

XX Methods for detecting target prostate-specific polynucleotides or  
PT diseases of the prostate (e.g. prostate cancer), comprising detecting  
PT the presence of any of PS108 nucleic acid sequences in a test sample -

XX Claim 1; Column 77-80; 55pp; English.

XX The present invention is related to a number of partial coding and  
CC protein sequences for the human prostate tissue protein PS108. These  
CC sequences can be used in the diagnosis and prognosis of prostate  
CC diseases, particularly prostate cancer. They can also be used to produce  
CC antibodies which can be used in treatment. The present sequence is one  
CC of the PS108 partial coding sequences.

XX Sequence 2133 BP; 414 A; 618 C; 561 G; 540 T; 0 other;

Query Match 56.2%; Score 1915; DB 21; Length 2133;  
Best Local Similarity 98.9%; Pred. No. 2.3e-290;  
Matches 2127; Conservative 0; Mismatches 5; Indels 19; Gaps 19;

QY 1177 GGGGCTGTACAGGGGCTGCCAGAGCTGAGCCGGGACCGAGGCGGAGACACTATCA 1236  
Db 1 GGGGCTGTACAGGGGCTGCCAGAGCTGAGCCGGGACCGAGGCGGAGACACTATCA 60

QY 1237 TGAAGGGTTTCGATGGGAGCCCTGGGGCTGTCTTCGAGTGGCCATCTCCCTGGTCTT 1296  
Db 61 TGAAGGGTTTCGATGGGAGCCCTGGGGCTGTCTTCGAGTGGCCATCTCCCTGGTCTT 120

QY 1297 CTCTCTGGTTCATGACCGGCTGTGTCAGGATTCGGCACTTCGAGCACTTATTTGGCCAG 1356  
Db 121 CTCTCTGGTTCATGACCGGCTGTGTCAGGATTCGGCACTTCGAGCACTTATTTGGCCAG 180

QY 1357 TGTGGCAGCTTTCCCTGTGGCTGCCGATGCCATGCCCTGTCACAGAGTGTGGCCGTGT 1416  
Db 181 TGTGGCAGCTTTCCCTGTGGCTGCCGATGCCATGCCCTGTCACAGAGTGTGGCCGTGT 240

QY 1417 GACAGCTTACGCGCCCTCACCGGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476  
Db 1417 GACAGCTTACGCGCCCTCACCGGGTTACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476



Db	241	GACAGCTTCAGCGCGCCTCACCGGGTTACCTTCTTCAGCCCTCGAGATCCTGCCCTACAC	300
Qy	1477	ACTGGCCTCCTCTACCAACCGGAGAGCAGGTGTTCTGTCGCCAAATACAGAGGGACAC	1536
Db	301	ACTGGCCTCCTCTACCAACCGGAGAGCAGGTGTTCTGTCGCCAAATACAGAGGGACAC	360
Qy	1537	TGGAGGTGCTAGCAGTAGAGACACGCTGTATGACCAAGCTTCCTGCCAGGCCCTAAGCCTGG	1596
Db	361	TGGAGGTGCTAGCAGTAGAGACACGCTGTATGACCAAGCTTCCTGCCAGGCCCTAAGCCTGG	420
Qy	1597	AGCTCCCTTCCCTAATGAGACACGTGGGTGCTGAGGACAGTGGCTGCTCCCACTCAACC	1656
Db	421	AGCTCCCTTCCCTAATGAGACACGTGGGTGCTGAGGACAGTGGCTGCTCCCACTCAACC	480
Qy	1657	CGCGCTCTCGGGGCCCTCTGCCGTGATGCTCCGTACGTGTGGTGTGGTGTGGTGTGGTGTGGT	1716
Db	481	CGCGCTCTCGGGGCCCTCTGCCGTGATGCTCCGTACGTGTGGTGTGGTGTGGTGTGGTGTGGT	540
Qy	1717	CGAGGCCAGGGTGGTTCGGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC	1776
Db	541	CGAGGCCAGGGTGGTTCGGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC	600
Qy	1777	CTTCTGCTCTCCAGGTGGCCCATCCCTGTTATATGGGCTCCATGTGCAGCTCAGCCA	1836
Db	601	CTTCTGCTCTCCAGGTGGCCCATCCCTGTTATATGGGCTCCATGTGCAGCTCAGCCA	660
Qy	1837	GTCTGCACCTGCCTATATGTTGTCTGCCGACAGCCTGGGTCTGGTGGCCATTTTACTTTGC	1896
Db	661	GTCTGCACCTGCCTATATGTTGTCTGCCGACAGCCTGGGTCTGGTGGCCATTTTACTTTGC	720
Qy	1897	TACACAGGTAGTATTTGACAAGACGCACTTGGGCCAAATACTCAGCCTAGAAAACTTCCAG	1956
Db	721	TACACAGGTAGTATTTGACAAGACGCACTTGGGCCAAATACTCAGCCTAGAAAACTTCCAG	780
Qy	1957	CACATTTGGGCTGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCTGCTCTGTTAGCCCCAT	2016
Db	781	CACATTTGGGCTGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCTGCTCTGTTAGCCCCAT	840
Qy	2017	GGGGCTGCCGGGCTGGCGCCAGTTTCTGTTGTGCCAAAGTAATGTGGCTCTCTGCTGC	2076
Db	841	GGGGCTGCCGGGCTGGCGCCAGTTTCTGTTGTGCCAAAGTAATGTGGCTCTCTGCTGC	900
Qy	2077	CACCTGTGCTGTAGGTGCGTAGTGCACAGCTGGGGCTGGGGCTGCCCTCTCCTCT	2136
Db	901	CACCTGTGCTGTAGGTGCGTAGTGCACAGCTGGGGCTGGGGCTGCCCTCTCCTCT	960
Qy	2137	CTCCCACTCTTAGGGCTGCCAGTGGAGGCTTCCAGAGGGGTTTCAGTCTGGACCTT	2196
Db	961	CTCCCACTCTTAGGGCTGCCAGTGGAGGCTTCCAGAGGGGTTTCAGTCTGGACCTT	1019
Qy	2197	ATACAGGGAGGCCAGAGGCTCCATGCATCGAATGCGGGGACTCTGCAAGGTGGATTAC	2256
Db	1020	ATACAGGGAGGCCAGAGGCTCCATGCATCGAATGCGGGGACTCTGCAAGGTGGATTAC	1078
Qy	2257	CGAGGCTCAGGGTTAACAGCTAGCTCCTTAGTTGAGACACACCTAGAGAAGGTTTTGG	2316
Db	1079	CGAGGCTCAGGGTTAACAGCTAGCTCCTTAGTTGAGACACACCTAGAGAAGGTTTTGG	1137
Qy	2317	GAGCTGAATAAATCTCAGTCACTGGTTTCCCATCTTAAGCCCCCTTAACCTGAGCTTCG	2376
Db	1138	GAGCTGAATAAATCTCAGTCACTGGTTTCCCATCTTAAGCCCCCTTAACCTGAGCTTC	1196
Qy	2377	TTTAAATGACTCTTGGCATGGGAGTTTCTAGGATGAACACTCCTCCATGGGATTTCG - AA	2435
Db	1197	TTTAAATGACTCTTGGCATGGGAGTTTCTAGGATGAACACTCCTCCATGGGATTTCGAAA	1256
Qy	2436	CATATGACTTATTGTAGGGGAAGTTCCTGAGGGGCAACACAGAACCGAGTCCCCCT	2495
Db	1257	TATGAAAGTTATTGTAGGGGAAGTTCCTGAGGGGCAACACAGAACCGAGTCCCCCT	1315
Qy	2496	CAGCCACAGACTGTCTTTTGGTGTATCACCCCCCTCTTACTTTTATCAGATGTGG	2555
Db	1316	CAGCCACAGACTGTCTTTTGGTGTATCACCCCCCTCTTACTTTTATCAGATGTGG - G	1374

Qy	2556	CCTGTTGGTCCCTCTCTGTTGCCATCACAGACACAGGCATTTAAATATTTAACTTATTTA	2615
Db	1375	CCTGTTGGTCCCTCTCTGTTGCCATCACAGACACAGGCATTTAAATATTTAACTTA-TTA	1433
Qy	2616	TTTAAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTTAATATTGGG	2675
Db	1434	TTTAAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTTAATATT-TT	1492
Qy	2676	TAGGTTGGGGGATCCCCAACAAATCAGTCCGCCCTGAGATAGCTGGTCATTTGGGCTGATCAT	2735
Db	1493	TAGGTTGGGGATCCCCAACAAATCAGTCCGCCCTGAGATAGCTGGTCATTTGGGCTGATC-T	1551
Qy	2736	TGCCAGAAATCTTCTTCCTCGGGGTCTGGGCCCCCAAAATGCTTAACCCAGGACCTTGGGA	2795
Db	1552	TGCCAGAAATCTTCTTCCTCGGGGTCTGGGCCCCCAAAATGCTTAACCCAGGACCTT-GA	1610
Qy	2796	AATTCTACTCATCCCAATCATTAATCCAAATGCTGTTACCCAAAGTTAGGGTGTGAAG	2855
Db	1611	AATTCTACTCATCCCAATCATTAATCCAAATGCTGTTACCCAAAGTTAGGGTGTGTG-AG	1669
Qy	2856	GAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTTAAGCACCCCTCTTCTCTTGG	2915
Db	1670	GAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTTAAGCACCCCTCTTCTCTT-G	1728
Qy	2916	CCAGCCTGGTTCCTCCGCCACTTCAGCTCCCTCTACTCTCTCTAGGACTGGGCTGATGAA	2975
Db	1729	CCAGCCTGGTTCCTCCGCCACTTCAGCTCCCTCTACTCTCTCTAGGACTGGGCTGATG-A	1787
Qy	2976	GGCACTGCCCAAAATTTCCCTCACCCTCACTTTCCCTCACCCTCACTTTCCCTCACTTTCC	3035
Db	1788	GGCACTGCCCAAAATTTCCCTCACCCTCACTTTCCCTCACCCTCACTTTCCCTCACTTTCC	1846
Qy	3036	CTCCACAACCTGTTTGGAGCTACTGCAGGACCAAGACAAAGTTCGGTTCCTCCCAAGC	3095
Db	1847	CTCCACAACCTGTTTGGAGCTACTGCAGGACCAAGACAAAGTTCGGTTCCTCCCAA-C	1905
Qy	3096	CTTTGTCCATCTCAGCCCCCAGAGTATCTGTGTTGGGGAATCTCACAGAAACTCA	3155
Db	1906	CTTTGTCCATCTCAGCCCCCAGAGTATCTGTGTTGGGGAATCTCACAGAAACT-A	1964
Qy	3156	GGAGCACCCCTGCCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAAAGTGCCGT	3215
Db	1965	GGAGCACCCCTGCCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAAAGTGCC-T	2023
Qy	3216	TTGCAATAATGTCGTCTTATTATTATTAGCGGGGTGAATATTTTATATCTGTAAGTGAGCAA	3275
Db	2024	TTGCAATAATGTCGTCTTATTATTATTAGCGGGGTGAATATTTTATATCTGTAAGTGAGC-A	2082
Qy	3276	TCAGAGTATATGTTTATGTGTGACAAATTTAAAGGCTTCTCTTATATGTTTA	3326
Db	2083	TCAGAGTATATGTTTATGTGTGACAAATTTAAAGGCTTCTCTTATATGTTTA	2133
RESULT 28			
AAS07155			
ID	AAS07155 standard; cDNA; 2133 BP.		
XX	AAS07155;		
XX			
DT	23-OCT-2001 (first entry)		
XX			
DE	Prostate gene PS108-specific cDNA (EST) consensus sequence.		
XX			
KW	Prostate; PS108; immunogen; drug screening; image localisation;		
KW	diagnostic; therapeutic; prostate tissue disease; cancer; metastasis;		
XX	expressed sequence tag; EST; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..799	
TT	/*tag= a		

RESULT 28

RESOL 2  
AAS07155

ID AAS07155 standard: cDNA: 2133 bp.

XX  
XX

AC AAS07155:

XX

DT 23-OCT-2001 (first entry)

XX  
XX  
XX

DE Prostate gene PS108-specific cDNA (EST) consensus sequence.

XX

KW Prostate; PS108; immunogen; drug screening; image localisation;

KW. diagnostic; therapeutic; prostate tissue disease; cancer; metastasis;

KW expressed sequence tag; EST; ss.

XX

OS Homo sapiens.

XX

FH	Key	Locati
----	-----	--------

FT	CDS	1..799
----	-----	--------

```
FT
/*tag= a
```





QY 2616 TTTAACAAGTAGAAGGAATCCATTGCTAGCTTTCTCTGTTGGTCTCTAATATTGGG 2675  
 DB 1434 TTTAACAAGTAGAAGGAATCCATTGCTAGCTTTCTCTGTTGGTCTCTAATATTGG 1492  
 QY 2676 TAGGGTGGGGATCCCAACAATCAGTCCCTGAGATAGCTGGTCAATGGGCTGATCAT 2735  
 DB 1493 TAGGGTGGGGATCCCAACAATCAGTCCCTGAGATAGCTGGTCAATGGGCTGATC-T 1551  
 QY 2736 TGGCAGAATCTTCTCTCTGGGCTGGCCGCCCAAAATGCTTAACCCAGGACCTTGA 2795  
 DB 1552 TGGCAGAATCTTCTCTCTGGGCTGGCCGCCCAAAATGCTTAACCCAGGACCTT-GA 1610  
 QY 2796 AATTCTACTCATCCCAAAATGATAATTCCAAATGCTTTACCAAGTCTAGGCTTGAAG 2855  
 DB 1611 AATTCTACTCATCCCAAAATGATAATTCCAAATGCTTTACCAAGTCTAGGCTTGAAG 1669  
 QY 2856 GAAGGTAGAGGTGGGGCTTCAGGCTCAACGGCTTCCCTAACCAACCCCTCTCTCTGG 2915  
 DB 1670 GAAGGTAGAGGTGGGGCTTCAGGCTCAACGGCTTCCCTAACCAACCCCTCTCTCTG 1728  
 QY 2916 CCCAGCCTGTTCCCGCCACTTCCACTCCCTCTACTCTCTAGGACTGGGCTGATGA 2975  
 DB 1729 CCCAGCCTGTTCCCGCCACTTCCACTCCCTCTACTCTCTAGGACTGGGCTGATG-A 1787  
 QY 2976 GGCACCTGCCCAAAATTTCCCTACCCCAACTTTCCCTTACCCCAACTTTCCCAACCA 3035  
 DB 1788 GGCACCTGCCCAAAATTTCCCTACCCCAACTTTCCCTTACCCCAACTTTCCCAACCA 1846  
 QY 3036 CTCACAAACCTGTTGGAGCTACTGCAGGACCAAGACAAAGTGGCGTTTCCCAAGC 3095  
 DB 1847 CTCACAAACCTGTTGGAGCTACTGCAGGACCAAGACAAAGTGGCGTTTCCCAAGC 1905  
 QY 3096 CTTTGTCCATCTCAGCCCCAGAGTATATCTGCTTGGGGAATCTCACACAAACTCA 3155  
 DB 1906 CTTTGTCCATCTCAGCCCCAGAGTATATCTGCTTGGGGAATCTCACACAAACTCA 1964  
 QY 3156 GGAGCACCCTGCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAAAGTCCCT 3215  
 DB 1965 GGAGCACCCTGCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAAAGTCCCT-T 2023  
 QY 3216 TTGCAATAATGCTCTTATTTATTTAGCGGGTGAATATTTATCTGTAAGTGAAGCA 3275  
 DB 2024 TTGCAATAATGCTCTTATTTATTTAGCGGGTGAATATTTATCTGTAAGTGAAGCA 2082  
 QY 3276 TCAGAGTATAATGTTTATGAGTACAAATTAAGGCTTCTTATATGTTTA 3326  
 DB 2083 TCAGAGTATAATGTTTATGAGTACAAATTAAGGCTTCTTATATGTTTA 2133

RESULT 29  
 AAC64927  
 ID AAC64927 standard; DNA; 2124 BP.  
 XX AAC64927;  
 AC  
 XX  
 XX  
 DT 07-FEB-2001 (first entry)  
 XX  
 DE Human prostate-related PS108 partial coding sequence SEQ ID NO: 15.  
 KW Human; prostate cancer; PS108; antibody; tumour; metastasis; ds.  
 XX Homo sapiens.  
 OS  
 XX  
 PN US6130043-A.  
 XX  
 PD 10-OCT-2000.  
 XX  
 PF 01-MAY-1998; 98US-0071710.  
 XX  
 PR 02-MAY-1997; 97US-0850713.  
 XX  
 PA (ABBO ) ABBOTT LAB.

XX Cohen M, Colpitts TL, Friedman PN, Gordon J, Granados EN;  
 PI Billing-Medel PA, Klass MR, Roberts-Rapp L, Stroupe SD, Yu H;  
 PI Kratochvil JB, Russell JC, Hodges SC;  
 XX WPI; 2000-655655/63.  
 XX Methods for detecting target prostate-specific polynucleotides or  
 PT diseases of the prostate (e.g. prostate cancer), comprising detecting  
 PT the presence of any of PS108 nucleic acid sequences in a test sample.  
 XX  
 XX Claim 1; Column 75-78; 55pp; English.  
 XX The present invention is related to a number of partial coding and  
 CC protein sequences for the human prostate tissue protein PS108. These  
 CC sequences can be used in the diagnosis and prognosis of prostate  
 CC diseases, particularly prostate cancer. They can also be used to produce  
 CC antibodies which can be used in treatment. The present sequence is one  
 CC of the PS108 partial coding sequences.  
 XX  
 XX Sequence 2124 BP; 411 A; 616 C; 558 G; 539 T; 0 other;  
 SQ  
 Query Match 55.7%; Score 1899.8; DB 21; Length 2124;  
 Best Local Similarity 99.0%; Pred. No. 5.5e-288;  
 Matches 2121; Conservative 0; Mismatches 2; Indels 20; Gaps 20;  
 QY 1185 ACCAGGGCGTGCACAGAGCTGAGCGGGCAGCGAGCGCCGAGACATATGATGAAGCG 1244  
 DB 1 ACCAGGGCGTGCACAGAGCTGAGCGGGCAGCGAGCGCCGAGACATATGATGAAGCG 60  
 QY 1245 TTCGATGGGAGCGCTGGGGTGTTCCTGAGTGGGCCATCTCCCTGCTCTCTCTGG 1304  
 DB 61 TTCGATGGGAGCGCTGGGGTGTTCCTGAGTGGGCCATCTCCCTGCTCTCTCTGG 120  
 QY 1305 TCATGGACCGCGCTGTGTCAGCATTTCGGCAGCTTCATTTGGCCAGTGGGCGAG 1364  
 DB 121 TCATGGACCGCGCTGTGTCAGCATTTCGGCAGCTTCATTTGGCCAGTGGGCGAG 180  
 QY 1365 CTTTCCCTGTGCTGCTGCGGTGCACATGCTCTCCACAGTGTGCGCCGTGTGACAGCTT 1424  
 DB 181 CTTTCCCTGTGCTGCTGCGGTGCACATGCTCTCCACAGTGTGCGCCGTGTGACAGCTT 240  
 QY 1425 CAGCGCCCTCTACCGGGTTCACCTTCTCAGCCCTCAGATCTGCCCTACACACTTGGCT 1484  
 DB 241 CAGCGCCCTCTACCGGGTTCACCTTCTCAGCCCTCAGATCTGCCCTACACACTTGGCT 300  
 QY 1485 CCTCTACACCGGGAGAGAGAGGTGTTCTCTCCCAATACCGAGGGGACACTGGAGGTG 1544  
 DB 301 CCTCTACACCGGGAGAGAGAGGTGTTCTCTCCCAATACCGAGGGGACACTGGAGGTG 360  
 QY 1545 CTAGAGTGGAGACAGCTGATGACAGCTTCTCTCCAGGCCCTAAGCCTGGAGTCCCT 1604  
 DB 361 CTAGAGTGGAGACAGCTGATGACAGCTTCTCTCCAGGCCCTAAGCCTGGAGTCCCT 420  
 QY 1605 TCCTTAATGGACAGTGGGTGTGAGGAGTGGCTGCTCCACCTCCACCCGGGCTCT 1664  
 DB 421 TCCTTAATGGACAGTGGGTGTGAGGAGTGGCTGCTCCACCTCCACCCGGGCTCT 480  
 QY 1665 GCGGGCCCTCTGCTGCTGATGCTCCGTACGTGTGGTGTGGGTGAGCCACCGAGGCCA 1724  
 DB 481 GCGGGCCCTCTGCTGCTGATGCTCCGTACGTGTGGTGTGGGTGAGCCACCGAGGCCA 540  
 QY 1725 GGGTGGTTCGGGCCCGGGGATCTGCTTGGACCTCGCCATCTGGATATGCTCTCTGTC 1784  
 DB 541 GGGTGGTTCGGGCCCGGGGATCTGCTTGGACCTCGCCATCTGGATATGCTCTCTGTC 600  
 QY 1785 TGTCCAGGTGGCCCATCCCTGTTTATGGCTCCATTTCCAGCTCAGCCAGTCTGTCA 1844  
 DB 601 TGTCCAGGTGGCCCATCCCTGTTTATGGCTCCATTTCCAGCTCAGCCAGTCTGTCA 660  
 QY 1845 CTGCTTATGTTGCTGCTGCGGAGGCTGGGTCTGCTGCCCATTTACTTTGTACACAGG 1904  
 DB 661 CTGCTTATGTTGCTGCTGCGGAGGCTGGGTCTGCTGCCCATTTACTTTGTACACAGG 720

Qy	1905	TAGTATTGACAAGAGCGACTTGGCCAAATACTCAGCTAGAAACTTTCAGACCATGG	1964
Db	721	TAGTATTGACAAGAGCGACTTGGCCAAATACTCAGCTAGAAACTTTCAGACCATGG	780
Qy	1965	GGTGGAGGGCCTGCCTCAGTGGTCCAGCTCCCGCTCCTGTTAGCCCCATGGGGTGC	2024
Db	781	GGTGGAGGGCCTGCCTCAGTGGTCCAGCTCCCGCTCCTGTTAGCCCCATGGGGTGC	840
Qy	2025	CGGGCTGGCCGCCAGTTTCTGTGCTGCCAAAGTAATGCGCTCCTGCTGCCACCTGT	2084
Db	841	CGGGCTGGCCGCCAGTTTCTGTGCTGCCAAAGTAATGCGCTCCTGCTGCCACCTGT	900
Qy	2085	GCTGCTGAGGTGGGTAGCTGCACAGCTGGGGGTGGGGCTCCCTCTCTCTCTCCCCAG	2144
Db	901	GCTGCTGAGGTGGGTAGCTGCACAGCTGGGGGTGGGGCTCCCTCTCTCTCTCCCCAG	960
Qy	2145	TCCTTAGGGCTGCCTGACTGGAGCGCTTCCAAGGGGTTTCAGTCTGGAGCTTATACAGGG	2204
Db	961	TCCTTAGGGCTGCCTGACTGGAGCGCTTCCAAGGGGTTTCAGTCTGGAGCTTATACAGGG	1019
Qy	2205	AGGCCAAGAGGGCTCCATGCACACTGGAATGCGGGGACTCTGCAGGTGGATACCCAGGCTC	2264
Db	1020	AGGCCAAGAGGGCTCCATGCACACTGGAATGCGGGGACTCTGCAGGTGGATACCCAGGCTC	1078
Qy	2265	AGGGTTTAACAGCTAGCGCTCCTAGTTGAGACACACTAGAGAAGGGTTTTTGGAGCTGAA	2324
Db	1079	-GGGTTTAACAGCTAGCGCTCCTAGTTGAGACACACTAGAGAAGGGTTTTTGGAGCTGAA	1137
Qy	2325	TAAACTCAGTCACCTGGTTTTCCCATCTCTAAGCCCTTAACTGACAGCTCGTTTTAATGT	2384
Db	1138	-AAACTCAGTCACCTGGTTTTCCCATCTCTAAGCCCTTAACTGACAGCTCGTTTTAATGT	1196
Qy	2385	AGCTCTTGATGGGAGTTCTAGGATGAACACTCTCCATGGGATTTGAACATAG - AC	2443
Db	1197	-GCTCTTGATGGGAGTTCTAGGATGAACACTCTCCATGGGATTTGAACATATGAAA	1255
Qy	2444	TTATTGTAGGGGAAGAGTCCTGAGGGGCAACACAGAACCAGCTCCCTCAGCCAC	2503
Db	1256	TTATTGTAGGGGAAGAGTCCTGAGGGGCAACACAGAACCAGCTCCCTCAGCCCA -	1314
Qy	2504	AGCACTGCTCTTTTGTCTGATCACCCCTCTTACGTTTTATCAGGATGGCGTGTGG	2563
Db	1315	AGCACTGCTCTTTTGTCTGATCACCCCTCTTACGTTTTATCAGGATGGCGTGT - G	1373
Qy	2564	TCCTCTCTGTGTCATCACAGACACAGGCATTTAAATATTTAACTATTATTAAACAA	2623
Db	1374	TCCTCTCTGTGTCATCACAGACACAGGCATTTAAATATTTAACTATTATTAAAC - A	1432
Qy	2624	AGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGCTTAATTTGGTAGGGTGG	2683
Db	1433	AGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGCTTAATTTGGTAGGGT - G	1491
Qy	2684	GGGATCCCCAACAAATCAGTGCCCTCGAGATAGCTGTCATTTGGGCTGATCTGCCAGAA	2743
Db	1492	GGGATCCCCAACAAATCAGTCCCCGAGATAGCTGTCATTTGGGCTGATCTGTCAG - A	1550
Qy	2744	TCCTTCTCTCTGCGGGCTGCGCCCCCCCCAAATGCCTAACCCAGGACCTTGGAAATCTAC	2803
Db	1551	TCCTTCTCTCTGCGGGCTGCGCCCCCCCCAAATGCCTAACCCAGGACCTTGGAAATCTA -	1609
Qy	2804	TCATCCCAAAATGATAATTCCTAAATGCTGTACCCAAAGGTTAGGGTGTGAAGGAAGTAG	2863
Db	1610	TCATCCCAAAATGATAATTCCTAAATGCTGTACCCAAAGGTTAGGGTGTGAAGGAAGTA -	1668
Qy	2864	AGGTTGGGGCTTCAGGTCTCAACGGCTTCCCTTAACCAACCCCTCTCTCTTGGCCCCAGCT	2923
Db	1669	AGGTTGGGGCTTCAGGTCTCAACGGCTTCCCTTAACCAACCCCTCTCTCTTGGCCCCAGCC -	1727
Qy	2924	GGTTCCTCCCCACCTTCCACTCCCTCTACTCTCTTAGGACTGGGCTGATGAAGCACTGC	2983
Db	1728	GGTTCCTCCCCACCTTCCACTCCCTCTACTCTCTTAGGACTGGGCTGATGAAGCACTGC -	1786

RESULT 30

AAS07601

ID AAS07601 standard; cDNA: 2124 BP.

AA  
AC AAS07601:XX  
DT 23-OCT-2001 (first entry)

Prostate gene PS108-specific

KW Prostate: D5109.

KW diagnostic; therapeutic; prostate tissue disease; cancer; metastasis;  
 KW expressed sequence tag; EST; ss.

XX  
XX  
OS  
Homo sapiens  
reference sequence tag, 451, 55.

OS  
XX  
Homo sapiens.

PN US6252047-B1.

PD 26-JUN-2001.

15-MAR-2000; PF

XX  
PR 01-MAY-1998. 000000-0000000000

PR 02-MAY-1997; 97US-0850713.

PA (ABBO ) ABBOTT LAB.

PI Billing-Medel PA,

PI Granados EN, Hodges SC, Kratochvil JD, Roberts-Rapp L;  
PI Russell JC, Stroupe SD, Yu H;  
PI

XX  
DR WPI; 2001-424488/45.

XX  
PT  
Naval Personnel  
WFL; 2001-424488/45.

Novel pS108 polypept

PS Example 1; Fig 1: 5500: English  
XX Example 1; Fig 1: 5500: English  
PS Example 1; Fig 1: 5500: English

example 1; Flg 1; 55pp; English.

The sequence represents the coding

was used along with other overlapping cDNA clones to produce a full length consensus sequence (see A507155). This sequence could then be used to produce pS108 polypeptides which are useful in assays for detecting antibodies to prostate tissue, and as immunogens to produce antibodies. The polypeptide is useful for screening compounds which specifically bind to the polypeptide and for screening for drugs, compounds, or any other agent which can be used to treat diseases associated with pS108. The antibody is useful to detect, or for image



Db 1905 ATCTAGCCCCCAGAGTATATCTGTGTGGGAATCTCACAGAACTCAGGAGCA-C 1963  
 QY 3164 CCCTGCTAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTAAGTGCCTTTGCAATA 3223  
 Db 1964 CCCTGCTAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTAAGTGCCTTTGCAAT- 2022  
 QY 3224 ATGTCGTCTATTTATTAGGGGGTGAATATTTATPACTGTAAAGTGAGCAATCAGAGTA 3283  
 Db 2023 ATGTCGTCTATTTATTAGGGGGTGAATATTTATPACTGTAAAGTGAGCAATCAGAGT- 2081  
 QY 3284 TAATGTTTATGTCACAAAATTAAGGCTTCTTATATGTTA 3326  
 Db 2082 TAATGTTTATGTCACAAAATTAAGGCTTCTTATATGTTA 2124

## RESULT 31

ABN81320  
 ID ABN81320 standard; cDNA; 3663 BP.

AC ABN81320;

DT 30-AUG-2002 (first entry)

XX Human mast cell related gene MC14 SEQ ID NO 3.

XX Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;  
 KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;  
 KW gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 294..746

FT /\*tag= a

FT /product= "MC14-1"

FT 1238..2218

FT /\*tag= a

FT /product= "MC14-2"

XX WO200246389-A2.

XX 13-JUN-2002.

XX 07-DEC-2001; 2001WO-US46180.

XX 08-DEC-2000; 2000US-251835P.

XX 14-MAR-2001; 2001US-275479P.

XX 28-MAR-2001; 2001US-279115P.

XX 02-APR-2001; 2001US-280143P.

XX (UNIO ) UCB SA.

XX Nocka K, Pirozzi G, Einstein R;

XX WPI; 2002-508560/54.

XX P-PSDB; ABB77570, ABB77571.

XX Novel isolated nucleic acids that are differentially expressed in mast

XX cells in patients with allergic hypersensitivity, encoding proteins

XX associated with mast cell degranulation and allergic hypersensitivity

XX

XX Claim 1; Page 98-102; 119pp; English.

XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),

XX corresponding to genes differentially expressed in mast cells following

XX activation or in patients with allergic hypersensitivity disease, (I)

XX that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of

XX (II) if at least 6 amino acids. (II) is useful for identifying binding

XX partners. (I) or (II) is useful for diagnosing or treating a disease

XX state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,

XX urticaria or atopic dermatitis or mastocytosis) in a subject which

XX involves determining the level of expression of (I) or (II). A computer

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC system, comprising a database containing information identifying the  
 CC expression level in a tissue or at least one mast cell of (I), is useful  
 CC for presenting information to identify the relative expression level of  
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic  
 CC response in a patient. The protein can also serve as a target that  
 CC modulate gene expression or activity and as an antigen to raise  
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying  
 CC agents that modulate expression of the protein or agents, such as  
 CC agonists or antagonists. The agonists or antagonists are useful for  
 CC modulating biological activity and function of (II) and thus are useful  
 CC for alleviating disease conditions such as allergic hypersensitivity,  
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.  
 CC XX  
 SQ Sequence 3663 BP; 740 A; 1058 C; 985 G; 880 T; 0 other;

Query Match 54.4%; Score 1854.8; DB 24; Length 3663;

Best Local Similarity 73.0%; Pred. No. 6.2e-281;

Matches 2667; Conservative 0; Mismatches 682; Indels 303; Gaps 10;

QY 2 GGAACACGCTGCAGCGCTGCTCCGGGTACACAGCCGCGCTCGGCCAGGATCTGAG 61  
 Db 12 GGAACACGCTGCAGCGCTGCTCCGGGTACACAGCCGCGCTCGGCCAGGATCTGAG 71  
 QY 62 TGATGAGACGTGTCTCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGCTGAGA 121  
 Db 72 TGATGAGACGTGTCTCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGCTGAGA 131  
 QY 122 AGCTGGACCGCCACCAAGGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGGAGTTG 181  
 Db 132 AGCTGGACCGCCACCAAGGGCTGGCAGAAATGGCGCTGGCTGATTCCTAGGAGTTG 191  
 QY 182 CCGCAGCAGGAGGAGGAGCGCGCAGCTTCTGGAGCAGAGCCGAGCAGCAGCTTCTGG 241  
 Db 192 CCGCAGCAGGAGGAGGAGCGCGCAGCTTCTGGAGCAGAGCCGAGCAGCAGCTTCTGG 251  
 QY 242 AGTGCCTGAACGGCCCTGAGCCCTACCGCCCTGGCCCACTATGTCAGAGGCTGTGG 301  
 Db 252 AGTGCCTGAACGGCCCTGAGCCCTACCGCCCTGGCCCACTATGTCAGAGGCTGTGG 311  
 QY 302 GTGAGCGCCCTGCTGCGGCACCGGAAAGCCAGCTTGTGTCACACCTGCTAACCTTT 361  
 Db 312 GTGAGCGCCCTGCTGCGGCACCGGAAAGCCAGCTTGTGTCACACCTGCTAACCTTT 371  
 QY 362 GGCCTGGAGGTGTGTTTGGCCGCGAGGCATCACCTATGTGCGGCTCTGCTGCTGGAAGTG 421  
 Db 372 GGCCTGGAGGTGTGTTTGGCCGCGAGGCATCACCTATGTGCGGCTCTGCTGCTGGAAGTG 431  
 QY 422 GGGGTAGAGGAGAGTTTCATGACCATTGCTGCTGGGAGGAGGCTGACCCCTGGGCTG 456  
 Db 432 GGGGTAGAGGAGAGTTTCATGACCATTGCTGCTGGGAGGAGGCTGACCCCTGGGCTG 491  
 QY 457 -----CAVTTGGTCCAGTGTGGGCTGGTGTGTGTCGCCG 491  
 Db 492 CTGTTCCAGATACATGCCACCTGGCAGTGTGGAGAGGAGTACCTCTGCCCTGGAGCTGC 551  
 QY 492 TCCTAGGCTCAGCAGTGACCACTGGGTGGAGCGCTATGCGCCGCCGCCCTTCATCT 551  
 Db 552 TTGGAGGAGAGGTGTCTGCTGGGAGGAGGCTGCTGGGAGGAGGCTGACCCCTGGGCTG 611  
 QY 552 GGG-----CACTGCTCTTGGGCAT 570  
 Db 612 AGGGGCGACACCAAGAGAAAGAGAAATACCAAGGACATACCCCACTGCTGGATC 671  
 QY 571 CTGTGCTGAGCCTCTTCTCATCCCAAGGCGCGCTGGCTAGCAGGCTGCTGTGCCCGGA 630  
 Db 672 CCTGTGCTCTCAGACAGAGCCTGGCTCATAGGAGACACTGGAGAAATGCTCTTAACCTTGG 731  
 QY 631 -----TCCAGAGCCCTGGAGCTG 649  
 Db 732 CTAGCCCTTTTATAATTTATAGGATATATCTCATTTAATGCTTACACCACTTTGAGG 791  
 QY 650 GCACTGCTCATCTCGGCGGTGGGCTGCTGGACTTCTGTGGCAGGCTGCTTCACTCCA 709

Db 792 TGATCCATTTTACAGAGAGGAGCAGAGGCTTTTAAAGAGTTAGTAACTTACGCCAA 851  
 Qy 710 CTGGAGGCTGCTCTCTGACCTCTTCGGGACCGGACCACTGTCGCCAGGCTACTCT 769  
 Db 852 AGCCAAATAGCAGCTGAACAGTAGAGCTGGAGCTCATCAAGGTCTCCCGCGGAGCTT 911  
 Qy 770 GTCTATGCTTTCATGATCAGTCTTG- 795  
 Db 912 GCTCCTACCCCTAGGACAAAGGGGTGGACTCTCTGACTCTGACAGTAAATTTTACAAAGCC 971  
 Qy 796 -----GGGCTGCTGGGCTACTCTCTGCTGCCATTTGACTGGGACCA 839  
 Db 972 ACAGAGGAAGTAGTAACATTTGTGACAAACCCCTACCCCGAGAGAGGGGCCCT 1031  
 Qy 840 GTCCCTGGCCCCCTACCTGGGACCCAGGAGAGTGCTCTTTTGGCCCTGCTCACCTCA 899  
 Db 1032 GTGAGATTTGACGGCTCTGGAGTCACACTGCTTTGTAACAGCTGCTCTTACCTCCCT 1091  
 Qy 900 TCTTCTCCTACGCTGAGGACCACTGCTGGTGGCTGAGGAGGAGCGCTGGGCCCCA 959  
 Db 1092 AGGTCTGGGCTTTGAATAGTATCACTTCTTAGTTGCTGCTCATGCTCATGTTGTCCATC 1151  
 Qy 960 ----- 959  
 Db 1152 TGAATATGGGGCATCTGTAATGCTGTGTTATGAGAGTAATATACAGCATCCCTGTGA 1211  
 Qy 960 -----CCGAGCCAGCAGAGGGCTGTCGGCC 985  
 Db 1212 AGAGTAGCACAGTCTGAGTAGGAGTATTTATTTCCATCTCTCACGAGACTTGGTTC 1271  
 Qy 986 CCTCTCTTTCGCCCCACTGCTGTCCATCCGCGGCCCTTGGCTTTCCGGAACCTGGGC 1045  
 Db 1272 CCTTCCCTTGGCCCTTACTTGTGCCAGCATTGACTCATACTACTTCCCTTCTTGCAG 1331  
 Qy 1046 GCCCTG-----CTTCCCGGCTGCACCACTGCTGCTGCCGATGCCCGCA 1091  
 Db 1332 GCATGGTCAGTGTGGGCTGGTCTGTGTCGCCCTCTAGGCTCAGCAGTACCACT 1391  
 Qy 1092 CCTTGGCGGCTCTTCTGCTGAGCTGTGACGCTGGATGGCAGCTGACATGACCTCACGC 1151  
 Db 1392 GCGTGGACGCTATGGCGCCGCCGCCCTTCTATCTGGGCACTGCTTGGGCACTCTGC 1451  
 Qy 1152 TGTTTTACAGGATTTGTGGGCGAGGGCTGTACAGGGGCTGTGCCAGAGCTGAGCGG 1211  
 Db 1452 TGAGGCTCTTCTCATCCCAAGGGCGGCTGGCTAGCAGGGCTGTGTGCCCGGATCCCA 1511  
 Qy 1212 GCACGAGCGCGGAGACACTATGATGAAGGGCTTCGGATGGCAGCCTGGGCTTCTCC 1271  
 Db 1512 GCGGCTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGTGGACTCTTGTGGCCAGG 1571  
 Qy 1272 TGCACTGCGC-CATCTCCCTGCTTCTCTCTGCTGATGGACCGGCTGTGACGAGATTC 1330  
 Db 1572 TGTGCTTCACTCCACTGGAGGCCCTGCTCTGACCTCTTCCGGACCGGACCACTGTC 1631  
 Qy 1331 GGCAGCTGAGCAGTCTATTTGGCAGTGTGGCAGCTTTCCTCTGCTGCTGCTGCCGAGCACA 1390  
 Db 1632 GCCAGGCTACTCTGTCTATGCTTATGATGATGCTTGGGCGTGGCTGCTGCTGCTGCTGCT 1691  
 Qy 1391 TGCTGTGCCACAGTGTGGCGTGTGA-----GAGCTTACGCGCCCTCACCGGGTTC 1444  
 Db 1692 TGCTGCTATGAGTGGGACACAGTGGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCT 1751  
 Qy 1445 ACCTTCTCAGCCCTGAGATCTCTGCTTACACACTGGGCTGCTTACCACTGGGAGAG 1504  
 Db 1752 GCTCTTTGGCTTGTCACTCCCTCATCTTCTCACTTGGCTAGCAGCCACACTGCTGGTG 1811  
 Qy 1505 CAGGTGCTTCTGCCC-----AAATACCGAGGGGACACTGGAGTGTCTAGCAGTGAAGACAG 1560  
 Db 1812 CTGAGGAGGAGCGCTGGGGCCCAACCGAGCCAGCAGAGGCTGTGGGCGGCTGCTGCTGCT 1871  
 Qy 1561 CTTGATACCACTTCTTGGCCAGGCTTACGCTTGGAGTCTGCTTCCCTCAATGGACAGCT 1620  
 Db 1872 CGCCCCACTGCTGTCCATGCCAGGCCCTAAGCCTGGAGTCTCCCTTCAATGGACAGCT 1931

Qy 1621 GGTGCTGGAGGAGTGGCTGCTCCACCTCCACCGCGCTCTGCGGGCTCTGCTG 1680  
 Db 1932 GGTGCTGGAGGAGTGGCTGCTCCACCTCCACCGCGCTCTGCGGGCTCTGCTG 1991  
 Qy 1681 TGATGCTCCCTAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740  
 Db 1992 TGATGCTCCCTAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2051  
 Qy 1741 GGGCATCTGCTGACCTGCGCATCTCTGGATAGTGGTGGTGGTGGTGGTGGTGGTGG 1800  
 Db 2052 GGGCATCTGCTGACCTGCGCATCTCTGGATAGTGGTGGTGGTGGTGGTGGTGGTGG 2111  
 Qy 1801 ATCCCTGTTTATGGCTCCATCTGCTCAGCTCAGCAGTCTCTCACTGCTATATGGTGT 1860  
 Db 2112 ATCCCTGTTTATGGCTCCATCTGCTCAGCTCAGCAGTCTCTCACTGCTATATGGTGT 2171  
 Qy 1861 TGCGCGAGGCTGCTGGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920  
 Db 2172 TGCGCGAGGCTGCTGGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2231  
 Qy 1921 CGACTTGGCCAAATACCTCAGCTAGAAACTTCCAGACATTTGGGTTGGGAGGCTGCT 1980  
 Db 2232 CGACTTGGCCAAATACCTCAGCTAGAAACTTCCAGACATTTGGGTTGGGAGGCTGCT 2291  
 Qy 1981 CACTGGTCTCCAGCTCCCGCTCTGTTAGCCCATGGGCTGCGGCTGCGCGCCAGT 2040  
 Db 2292 CACTGGTCTCCAGCTCCCGCTCTGTTAGCCCATGGGCTGCGGCTGCGCGCCAGT 2351  
 Qy 2041 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
 Db 2352 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2411  
 Qy 2101 GCTGACAGCTGGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160  
 Db 2412 GCTGACAGCTGGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2471  
 Qy 2161 ACTGAGGCTTCCAAAGGGGTTTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220  
 Db 2472 ACTGAGGCTTCCAAAGGGGTTTTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2531  
 Qy 2221 ATGCACTGGAATGGGAGTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280  
 Db 2532 ATGCACTGGAATGGGAGTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2591  
 Qy 2281 CTCTAGTGTAGACACCTAGAGAGGTTTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340  
 Db 2592 CTCTAGTGTAGACACCTAGAGAGGTTTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2651  
 Qy 2341 GTTTCCTCATCTTAAGCCCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
 Db 2652 GTTTCCTCATCTTAAGCCCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2711  
 Qy 2401 TTTCTAGATGAACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2458  
 Db 2712 TTTCTAGATGAACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2771  
 Qy 2459 GAGTCTGAGGGGCAACACAGAGGTTTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2518  
 Db 2772 GAGTCTGAGGGGCAACACAGAGGTTTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2831  
 Qy 2519 GTGATCCACCCCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2578  
 Db 2832 GTGATCCACCCCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2891  
 Qy 2579 CACAGACACAGGCTTAAATATTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2638  
 Db 2892 CACAGACACAGGCTTAAATATTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2951  
 Qy 2639 ATTGCTAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2698  
 Db 2952 ATTGCTAGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3011







2010	QY	GC	CCCATGGGGCTGCGGGCTGCGCGCAGTTCTTGTGCTGCTGCAAGTAATGTGGCTCT	2069
5652	DB	GC	CCCATGGGGCTGCGGGCTGCGCGCAGTTCTTGTGCTGCTGCAAGTAATGTGGCTCT	5711
2070	QY	CT	GCTGCACCCCTGTGCTGCTGAGTGTGCGTAGCTGCACAGCTGCGGGCTGGGGCTCCCT	2129
5712	DB	CT	GCTGCACCCCTGTGCTGCTGAGTGTGCGTAGCTGCACAGCTGGGGCTGGGGCTCCCT	5771
2130	QY	CT	CCTCTCTCTCCCGAGTCTCTAGGGCTGCCTTGACTGGAGGCTTCCAAAGGGGGTTTCAGTC	2189
5772	DB	CT	CCTCTCTCTCCCGAGTCTCTAGGGCTGCCTTGACTGGAGGCTTCCAAAGGGGGTTTCAGTC	5831
2190	QY	TG	GACTTATACAGGGAGGCCAGAAAGGCTCCATGCTGCACTGGAATGCGGGGACTCTGAGGT	2249
5832	DB	TG	GACTTATACAGGGAGGCCAGAAAGGCTCCATGCTGCACTGGAATGCGGGGACTCTGAGGT	5891
2250	QY	GG	ATTACCAGGCTCAGGGTTAAACAGCTAGCCCTCCTAGTTGACACACACTAGAGAAGGG	2309
5892	DB	GG	ATTACCAGGCTCAGGGTTAAACAGCTAGCCCTCCTAGTTGACACACACTAGAGAAGGG	5951
2310	QY	TTTT	TTGGGAGCTGAATAAATCACTAGTCACCTTGGTTTCCCATCTCTAAAGCCCTTAAACCTGC	2369
5952	DB	TTTT	TTGGGAGCTGAATAAATCACTAGTCACCTTGGTTTCCCATCTCTAAAGCCCTTAAACCTGC	6011
2370	QY	AG	CTTGCTTAAATGTAGCTCTCCATGGGAGTTTCTAGGATGAAACACTCTCCATGGGA	2429
6012	DB	AG	CTTGCTTAAATGTAGCTCTCCATGGGAGTTTCTAGGATGAAACACTCTCCATGGGA	6071
2430	QY	TTTT	GAACATATGAC--TTATTCTAGGGGAAGAGTCTGAGGGGCACACACAGAARCCA	2487
6072	DB	TTTT	GAACATATGAAAGTTATTCTAGGGGAAGAGTCTGAGGGGCACACACAGAARCCA	6131
2488	QY	GG	TCCTCCTCAGCCCCAGCACACTCTCTTTTGTCTGATCCACCCCTCTTACCTTTTATCA	2547
6132	DB	GG	TCCTCCTCAGCCCCAGCACACTCTCTTTTGTCTGATCCACCCCTCTTACCTTTTATCA	6191
2548	QY	GG	ATGTGGCTGTGGTCTCTCTGTTGCCATCACAGAGACACAGGCAATTTAAATATTTAA	2607
6192	DB	GG	ATGTGGCTGTGGTCTCTCTGTTGCCATCACAGAGACACAGGCAATTTAAATATTTAA	6251
2608	QY	CT	TATTATTAAACAAAGTAGAAGGAATCCATTGCTAGCTTTCTGTGTTGCTGCTAA	2667
6252	DB	CT	TATTATTAAACAAAGTAGAAGGAATCCATTGCTAGCTTTCTGTGTTGCTGCTAA	6311
2668	QY	TATT	GGTGAAGGTGGGGATCCCAACATCAGTCCCTGAGTAGTGGTTCATTTGGG	2727
6312	DB	TATT	GGTGAAGGTGGGGATCCCAACATCAGTCCCTGAGTAGTGGTTCATTTGGG	6371
2728	QY	CT	GATCATTTGCCAGAAATCTTCTCTGTTGGGCTGCGCCCCCAAAATGCTTAACCCAGG	2787
6372	DB	CT	GATCATTTGCCAGAAATCTTCTCTGTTGGGCTGCGCCCCCAAAATGCTTAACCCAGG	6431
2788	QY	AC	CTTGGAAATCTACTCATCCCAATGATAATTCAAATGCTTTACCCAAGTTAGGG	2847
6432	DB	AC	CTTGGAAATCTACTCATCCCAATGATAATTCAAATGCTTTACCCAAGTTAGGG	6491
2848	QY	TG	TGAAGGAAGGTAGAGGTTGGGCTTCAGGTCTCAAGGCTTCCTCAACCAACCCCTCT	2907
6492	DB	TG	TGAAGGAAGGTAGAGGTTGGGCTTCAGGTCTCAAGGCTTCCTCAACCAACCCCTCT	6551
2908	QY	TC	TTTGGGCCAGCTGTGGTTCCTCCCTTCCACTCCCTCTACTCTCTAGGACTGGG	2967
6552	DB	TC	TTTGGGCCAGCTGTGGTTCCTCCCTTCCACTCCCTCTACTCTCTAGGACTGGG	6611
2968	QY	CT	GATGAGCACTGCCCAAAATTTCCCTTACCCCAACTTTCCCTACCCCAACTTTC	3027
6612	DB	CT	GATGAGCACTGCCCAAAATTTCCCTTACCCCAACTTTCCCTACCCCAACTTTC	6671
3028	QY	CC	CACAGCTTCCACAAACCTGTTTGGAGTCTACTGCGAGGACGAGAAGCAAAAGTCGGTT	3087
6672	DB	CC	CACAGCTTCCACAAACCTGTTTGGAGTCTACTGCGAGGACGAGAAGCAAAAGTCGGTT	6731
3088	QY	TCCC	AAGCCCTTTGCTCCATCTCAGCCCCCAGAGTATATCTGCTTGGGAAATCTCACACA	3147

Db		6732	TCCCAAGCCTTGTGTCATCTCAGCCCCAGAGTATATCTGCTTGSGGAATCTCACACA	6799
QY		3148	GAAACTCAGGAGACACCCCCTGCCTGAGCTAAGGAGGCTCTATNCTCTCAGGGGGGGTTTA	3207
Dd		6792	GAAACTCAGGAGACACCCCCTGCCTGAGCTAAGGAGGCTCTATNCTCTCAGGGGGGGTTTA	6851
QY		3208	AGTGCCCGTTTCCAATAATGTCGCTTATTATTATTTAGCGGGGTGAATATTTTATACTGTAA	3267
Dd		6852	AGTGCCCGTTTCCAATAATGTCGCTTATTATTATTTAGCGGGGTGAATATTTTATACTGTAA	6911
QY		3268	GTGAGCAATCAGAGATATAATGTTTATGGTGACAAAATAAAGGCTTCTTATATGTTTAA	3327
Dd		6912	GTGAGCAATCAGAGATATAATGTTTATGGTGACAAAATAAAGGCTTCTTATATGTTTAA	6971
QY		3328	AAAAA 3332	
Dd		6972	AAAAA 6976	
 RESULT 33 AAH93869				
ID		AAH93869	standard; cDNA; 6976 BP.	
XX		AC		
XX		AAH93869;		
DT		04-OCT-2001	(first entry)	
XX				
DE		P553S	cDNA splice variant P553S-6.	
XX				
KW		Human;	prostate cancer; prostate-specific; diagnosis; vaccine;	
KW		cystatic;	gene therapy; metastasis; ss.	
OS		Homo sapiens.		
XX				
PN		WO200151633-A2.		
PD		19-JUL-2001.		
XX				
PF		16-JAN-2001;	2001WO-US01574.	
XX				
PR		14-JAN-2000;	2000US-0483672.	
XX				
PA		(CORI-) CORIXA CORP.		
PI		Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;		
PI		Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;		
PI		Wang A, Meagher MJ;		
XX				
DR		WPI; 2001-425873/45.		
XX				
PT		New polynucleotide encoding a prostate-specific protein, for		
PT		diagnosing, monitoring and treating prostate cancer in a patient and		
PT		for use in vaccines -		
XX				
PS		Claim 1; Page 461-463; 543pp; English.		
XX				
CC		The present invention describes polynucleotide sequences (I) which encode		
CC		prostate-specific proteins (II). (I) and (II) have cytostatic activity,		
CC		and can be used in vaccine production and gene therapy. (I), (II),		
CC		antibodies to (II), fusion proteins comprising (II), and isolated		
CC		T cells prepared using (I) or (II) are used treat cancer in a patient.		
CC		(I) and the antibodies are also used in the detection of cancer in a		
CC		patient. The cancer that is diagnosed or treated is particularly		
CC		prostate cancer. (I) and (II) can be used in vaccines. The antibodies or		
CC		(I) can be used for monitoring the progression of cancer in a patient.		
CC		(I) and (II) can also be used to improve diagnostic and therapeutic		
CC		methods for prostate cancer. They can indicate the level of metastasis		
CC		as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to		
CC		AAH01118 represent polynucleotide and amino acid sequences used in the		
CC		exemplification of the present invention.		
XX				
SO		Sequence 6976 BP; 1312 A; 1996 C; 1961 G; 1706 T; 1 other:		

Query Match 53.2%; Score 1815.8; DB 22; Length 6976; Best Local Similarity 98.7%; Pred. No. 8.2e-275; Matches 1841; Conservative 0; Mismatches 22; Indels 2; Gaps 1;			
QY	1470	CCTACACACTGGCTCCCTCTACACACCGGAGAGAGAGTGTCTCCGCCCAATACCGAG	1529
Db	5112	CCTGCTCTTCCCTTTCTACACCCCTCTGCTTAGGTGTCTCTGCCCAATACCGAG	5171
QY	1530	GGGACACTGGAGGTGCTAGCAGTGGAGCAGCAGCTGATGACCACTTCTCCAGGCGCTA	1589
Db	5172	GGGACACTGGAGGTGCTAGCAGTGGAGCAGCAGCTGATGACCACTTCTCCAGGCGCTA	5231
QY	1590	AGCCTGGAGCTCCCTCCCTAATGAGACAGCTGGGTGCTGGAGCAGTGGCTCTCCAC	1649
Db	5232	AGCCTGGAGCTCCCTCCCTAATGAGACAGCTGGGTGCTGGAGCAGTGGCTCTCCAC	5291
QY	1650	CTCCACCGGCTCTGGGGGCTCTGCTGTGATGCTCCGTACGTGGTGGTGGG	1709
Db	5292	CTCCACCGGCTCTGGGGGCTCTGCTGTGATGCTCCGTACGTGGTGGTGGG	5351
QY	1710	AGCCACCGAGGCTGGGTCCGGGCGGGGCTCTGCTGGAGCTCGCCATCTGG	1769
Db	5352	AGCCACCGAGGCTGGGTCCGGGCGGGGCTCTGCTGGAGCTCGCCATCTGG	5411
QY	1770	ATAGTGCCTTCTGCTGCCAGGTGGCCCATCCCTGTTTATGGGCTCCATGTCAGC	1829
Db	5412	ATAGTGCCTTCTGCTGCCAGGTGGCCCATCCCTGTTTATGGGCTCCATGTCAGC	5471
QY	1830	TCAGCAGTCTGCTACCTGCTATATGCTGCTGCGGAGGCTGGGTCTGGCCATTT	1889
Db	5472	TCAGCAGTCTGCTACCTGCTATATGCTGCTGCGGAGGCTGGGTCTGGCCATTT	5531
QY	1890	ACTTGTGTACACAGTATGTTTACACAGAGCGACTTGGCCAAATACTACGCTAGAAA	1949
Db	5532	ACTTGTGTACACAGTATGTTTACACAGAGCGACTTGGCCAAATACTACGCTAGAAA	5591
QY	1950	CTTCAGCACATGGGTGGAGGCTGCTCAGTGGGTGCCAGCTCCCGCTCTGTTA	2009
Db	5592	CTTCAGCACATGGGTGGAGGCTGCTCAGTGGGTGCCAGCTCCCGCTCTGTTA	5651
QY	2010	GCCCATGGGCTGCCGGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2069
Db	5652	GCCCATGGGCTGCCGGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5711
QY	2070	CTGCTGCCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2129
Db	5712	CTGCTGCCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5771
QY	2130	CTCCTCTCTCCAGTCTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2189
Db	5772	CTCCTCTCTCCAGTCTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5831
QY	2190	TGGACTTATACAGGAGGAGGAGGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT	2249
Db	5832	TGGACTTATACAGGAGGAGGAGGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT	5891
QY	2250	GGATTACCGAGCTCAGGCTTACAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2309
Db	5892	GGATTACCGAGCTCAGGCTTACAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5951
QY	2310	TTTTTGGAGCTGAATAAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2369
Db	5952	TTTTTGGAGCTGAATAAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6011
QY	2370	AGCTTCGTTTATGTAGCTCTTGGATGGAGTCTTCTAGGATGAACACTCTCCATGGGA	2429
Db	6012	AGCTTCGTTTATGTAGCTCTTGGATGGAGTCTTCTAGGATGAACACTCTCCATGGGA	6071
QY	2430	TTTGAAATATGAC--TTATTGTAGGGAAGAGTCTGTAGGGGCAACACACAGAACC	2487
Db	6072	TTTGAAATATGAAAGTTATTGTAGGGAAGAGTCTGTAGGGGCAACACACAGAACC	6131
QY	2488	GGTCCCTCAGCCACAGCACTGTCTTTTGTGTATGATCCACCCCTCTTACCTTTTATCA	2547
Db	6132	GGTCCCTCAGCCACAGCACTGTCTTTTGTGTATGATCCACCCCTCTTACCTTTTATCA	6191
QY	2548	GGATGGGCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2607
Db	6192	GGATGGGCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6251
QY	2608	CTTATTATTATTAACAAAGTAGAAGGAATCCATTTCTAGCTTTTCTGTTGGTCTCTAA	2667
Db	6252	CTTATTATTATTAACAAAGTAGAAGGAATCCATTTCTAGCTTTTCTGTTGGTCTCTAA	6311
QY	2668	TATTTGGGTAGGTGGGGATCCCAACAATCAGGTCCCTCAGATAGTGTGCTATGGG	2727
Db	6312	TATTTGGGTAGGTGGGGATCCCAACAATCAGGTCCCTCAGATAGTGTGCTATGGG	6371
QY	2728	CTGATCATTTGCCAGAAATCTTCTTCTCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT	2787
Db	6372	CTGATCATTTGCCAGAAATCTTCTTCTCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT	6431
QY	2788	ACCTTGGAAATTTCTACTCATCCCAAAATGATAATTCCTGCTGCTGCTGCTGCTGCTGCT	2847
Db	6432	ACCTTGGAAATTTCTACTCATCCCAAAATGATAATTCCTGCTGCTGCTGCTGCTGCTGCT	6491
QY	2848	TGTTGAAGAAAGTAGAGGTGGGCTTCAGTCTCAAGGGCTTCCCTAACCCCTCT	2907
Db	6492	TGTTGAAGAAAGTAGAGGTGGGCTTCAGTCTCAAGGGCTTCCCTAACCCCTCT	6551
QY	2908	TCCTTGGCCAGCCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2967
Db	6552	TCCTTGGCCAGCCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6611
QY	2968	CTGATGAAGGCACTGCCCAAAATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3027
Db	6612	CTGATGAAGGCACTGCCCAAAATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6671
QY	3028	CCCAAGCTCCACACACCTGTTTGGAGCTACTGAGGACCAAGAGCAAGAGTGGGTT	3087
Db	6672	CCCAAGCTCCACACACCTGTTTGGAGCTACTGAGGACCAAGAGCAAGAGTGGGTT	6731
QY	3088	TCCCAAGCTTTCTCCATCTCAGCCCTCAGATATATCTGCTGCTGCTGCTGCTGCTGCTGCT	3147
Db	6732	TCCCAAGCTTTCTCCATCTCAGCCCTCAGATATATCTGCTGCTGCTGCTGCTGCTGCTGCT	6791
QY	3148	GAACTCAGGAGCACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3207
Db	6792	GAACTCAGGAGCACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6851
QY	3208	AGTGGGCTTTGCAATTAATGCTGCTTATTTTATTTAGCGGGTGAATATTTATCTGTA	3267
Db	6852	AGTGGGCTTTGCAATTAATGCTGCTTATTTTATTTAGCGGGTGAATATTTATCTGTA	6911
QY	3268	GTGAGCAATCAGAGTATAATGTTTATGTTGACAAATTAAGGCTTCTTATATGTTAA	3327
Db	6912	GTGAGCAATCAGAGTATAATGTTTATGTTGACAAATTAAGGCTTCTTATATGTTAA	6971
QY	3328	AAAAA 3332	
Db	6972	AAAAA 6976	
RESULT 34			
ABL95412			
ID	ABL95412	standard; cDNA; 6976 bp.	
XX	ABL95412;		
XX	AC		
XX	AC		
DT	19-JUL-2002	(first entry)	
XX	Human P553S	splice variant SEQ ID NO 705.	
DE	Human;	cancer; prostate cancer; vaccine; cytostatic; immunostimulant;	
XX	Human;	cancer; prostate cancer; vaccine; cytostatic; immunostimulant;	
KW	gene therapy; gene; ss.		

XX	Homo sapiens.	1470	CCTACACTGGCTCCCTCTACACCGGAGAGCAGGTGTTCTCTGCCCAATACCGAG	1529
OS				
XX	US2002022248-A1.	5112	CTGTCTCTTCCCTTTCTTCAACCCCTCTGCCTTAGTGTCTCTGCCCAATACCGAG	5171
PN				
XX	21-FEB-2002.	1530	GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACACAGTCTCTGCCAGGCCCTA	1589
PD				
XX		5172	GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACACAGTCTCTGCCAGGCCCTA	5231
PF				
XX	12-JAN-2001; 2001US-0759143.	1590	AGCCTGGAGCTCCCTTCCCTTAATGGACACCTGGGTGCTGAGGACAGTGGCTGCTCCAC	1649
XX	25-FEB-1997; 97US-0806099.			
PR	01-AUG-1997; 97US-0904804.	5232	AGCCTGGAGCTCCCTTCCCTTAATGGACACCTGGGTGCTGAGGACAGTGGCTGCTCCAC	5291
PR	09-FEB-1998; 98US-0020956.			
PR	25-FEB-1998; 98US-0030607.	1650	CTCCACCCGGGCTCTGCGGGGCTCTGCTGCTGATGCTCCGACAGTGTGGTGGGTG	1709
PR	14-JUL-1998; 98US-0115453.			
PR	23-SEP-1998; 98US-0159812.	5292	CTCCACCCGGGCTCTGCGGGGCTCTGCTGCTGATGCTCCGACAGTGTGGTGGGTG	5351
PR	15-JAN-1999; 99US-0232149.			
PR	09-APR-1999; 99US-0288946.	1710	AGCCACCCGAGGCTGGTTCGGGGCGGGGCTGCTGCTGAGGACCTGCCATCCTGG	1769
PR	13-JUL-1999; 99US-0352616.			
PR	12-NOV-1999; 99US-0439313.	5352	AGCCACCCGAGGCTGGTTCGGGGCGGGGCTGCTGCTGAGGACCTGCCATCCTGG	5411
PR	18-NOV-1999; 99US-0443686.			
PR	14-JAN-2000; 2000US-0483672.	1770	ATAGTGGCTTCTGCTGCTCCAGGTGGCCCATCCCTGTTTATGGGCTCCATTTGCCAGC	1829
PR	27-MAR-2000; 2000US-0536857.			
PR	09-MAY-2000; 2000US-0568100.	5412	ATAGTGGCTTCTGCTGCTCCAGGTGGCCCATCCCTGTTTATGGGCTCCATTTGCCAGC	5471
PR	12-MAY-2000; 2000US-0570737.			
PR	13-JUN-2000; 2000US-0593793.	1830	TCAGCCAGTCTGTGCTGCTGCTATATGCTGCTGCGCAGGCTGGGTCTGGTGGCCATTT	1889
PR	27-JUN-2000; 2000US-0605783.			
PR	10-AUG-2000; 2000US-0636215.	5472	TCAGCCAGTCTGTGCTGCTGCTATATGCTGCTGCGCAGGCTGGGTCTGGTGGCCATTT	5531
PR	29-AUG-2000; 2000US-0651236.			
PR	06-SEP-2000; 2000US-0657279.	1890	ACTTTGCTACACAGGTAGTATTTCACAGAGGAGCTTGGCCCAATATCTCAGCGTAGAAA	1949
PR	02-OCT-2000; 2000US-0679426.			
PR	10-OCT-2000; 2000US-0685166.	5532	ACTTTGCTACACAGGTAGTATTTCACAGAGGAGCTTGGCCCAATATCTCAGCGTAGAAA	5591
XX	(XULJ/) XU J.			
PA	(DILL/) DILLON D C.	1950	CTTCCAGCACATTTGGGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2009
PA	(MITC/) MITCHAM J L.			
PA	(HARL/) HARLOCKER S L.	5592	CTTCCAGCACATTTGGGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5651
PA	(JIAN/) JIANG Y.			
PA	(KALO/) KALOS M D.	2010	CCCCATGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCT	2069
PA	(FANG/) FANGER G R.			
PA	(RETT/) RETTER W W.	5652	CCCCATGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCTGCGGGCT	5711
PA	(STOL/) STOLK J A.			
PA	(DAYC/) DAY C H.	2070	CTGCTGCCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2129
PA	(VEDV/) VEDVICK T S.			
PA	(CART/) CARTER D.	5712	CTGCTGCCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5771
PA	(LISX/) LI S X.			
PA	(WANG/) WANG A.	2130	CTCCTCTCTCCAGTCTCTAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2189
PA	(SKEI/) SKEIKY Y A W.			
PA	(HEPL/) HEPLER W T.	5772	CTCCTCTCTCCAGTCTCTAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5831
PA	(HEND/) HENDERSON R A.			
XX	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;			
PI	Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;			
PI	Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;			
XX	WPI: 2002-255649/30.			
XX				
XX	New prostate-specific polynucleotides for diagnosing and treating			
PT	diseases, in particular prostate cancer, and as markers for the			
PT	progression of cancer			
XX	Claim 1; SEQ ID NO 705; 87pp; English.			
XX				
CC	The present invention provides prostate-specific coding sequences and			
CC	their encoded proteins. These can be used in the diagnosis and treatment			
CC	of cancers, particularly prostate cancer. The present sequence is a cDNA			
CC	described in the invention.			
XX	Sequence 6976 BP; 1312 A; 1996 C; 1961 G; 1706 T; 1 other;			
SQ				
Query Match	53.2%; Score 1815.8; DB 24; Length 6976;			
Best Local Similarity	98.7%; Pred. No. 8.2e-275;			
Matches 1841; Conservative	0; Mismatches 22; Indels 1;			

```

Db 6192 GAGTGGGCTGTGGTCTCTCTGTTGCCATCACAGACACAGGATTAAATATTAA 6251
Qy 2608 CTTATTATTATTAACAAGTAGAAGGAATCCATTGCTAGCTTTCTGTTGTTGCTCTAA 2667
Db 6252 CTTATTATTATTAACAAGTAGAAGGAATCCATTGCTAGCTTTCTGTTGTTGCTCTAA 6311
Qy 2668 TATTGGGTAGGTGGGGGATCCCAACAATCAGTCCCTCAGATAGTGTGTCATTGGG 2727
Db 6312 TATTGGGTAGGTGGGGGATCCCAACAATCAGTCCCTCAGATAGTGTGTCATTGGG 6371
Qy 2728 CTGATCATTCAGAGATCTCTCTCCGGGCTGCGCCCAAAATGCCCTAACCCAGG 2787
Db 6372 CTGATCATTCAGAGATCTCTCTCCGGGCTGCGCCCAAAATGCCCTAACCCAGG 6431
Qy 2788 ACCTGGAAATTTACTATCATCCAAATGATTAATCCAAATGCTTTACCAAGGTTAGGG 2847
Db 6432 ACCTGGAAATTTACTATCATCCAAATGATTAATCCAAATGCTTTACCAAGGTTAGGG 6491
Qy 2848 TGTGGAAGGAGTAGAGGTGGGCTTCAGGTCTCAAGGGTTCCTTAACCAACCCCTCT 2907
Db 6492 TGTGGAAGGAGTAGAGGTGGGCTTCAGGTCTCAAGGGTTCCTTAACCAACCCCTCT 6551
Qy 2908 TCTCTGGCCAGCTGTGTTCCCGCCACTTCCACCTCCCTCTACTCTCTAGGACTGG 2967
Db 6552 TCTCTGGCCAGCTGTGTTCCCGCCACTTCCACCTCCCTCTACTCTCTAGGACTGG 6611
Qy 2968 CTGATGAAGGCACTGCCCAAAATTTCCCTACCCCAACTTTCCCTACCCCAACTTTC 3027
Db 6612 CTGATGAAGGCACTGCCCAAAATTTCCCTACCCCAACTTTCCCTACCCCAACTTTC 6671
Qy 3028 CCCACAGCTCCACACACCTGTTTGGAGCTACTGAGCAGCAGACAGACAAAGTCCGGTT 3087
Db 6672 CCCACAGCTCCACACACCTGTTTGGAGCTACTGAGCAGCAGACAGACAAAGTCCGGTT 6731
Qy 3088 TCCCAAGCTTTGTCTATCTCAGCCCCCAGAGATATATCTGTGCTTGGGGAATCTCACACA 3147
Db 6732 TCCCAAGCTTTGTCTATCTCAGCCCCCAGAGATATATCTGTGCTTGGGGAATCTCACACA 6791
Qy 3148 GAACTCAGGACACCCCTGCTCAGTACAGTACAGGAGGCTTATCTCTCAGGGGGGTTTA 3207
Db 6792 GAACTCAGGACACCCCTGCTCAGTACAGTACAGGAGGCTTATCTCTCAGGGGGGTTTA 6851
Qy 3208 AGTGGCTTTGCAATATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATCTGTAA 3267
Db 6852 AGTGGCTTTGCAATATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATCTGTAA 6911
Qy 3268 GTGAGCAATCAGAGTAAATGTTTATGTGACAAAATTAAGGCTTTCTTATATGTTTAA 3327
Db 6912 GTGAGCAATCAGAGTAAATGTTTATGTGACAAAATTAAGGCTTTCTTATATGTTTAA 6971
Qy 3328 AAAAA 3332
Db 6972 AAAAA 6976

```

RESULT 35

ABN81324

ID ABN81324 standard; cDNA: 1662 BP.

XX AC ABN81324;

30-AUG-2002 (first entry)

Human mast cell related splice variant gene MC14 SEQ ID NO 12.

Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;  
 KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;  
 gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

```

CDS 1..1662
FT FT /*tag= a
FT FT /product= "MC14 alternatively spliced variant"
PN WO200246389-A2.
XX 13-JUN-2002.
XX 07-DEC-2001; 2001WO-US46180.
XX 08-DEC-2000; 2000US-251835P.
XX 14-MAR-2001; 2001US-279479P.
XX 28-MAR-2001; 2001US-279115P.
XX 02-APR-2001; 2001US-280143P.
XX (UNIO ) UCB SA.
XX Nocka K, Pirozzi G, Einstein R;
XX WPI; 2002-508560/54.
XX P-PSDB; ABB77575.
XX Novel isolated nucleic acids that are differentially expressed in mast
XX cells in patients with allergic hypersensitivity, encoding proteins
XX associated with mast cell degranulation and allergic hypersensitivity
XX
XX Claim 1; Page 115-117; 119pp; English.
XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),
XX corresponding to genes differentially expressed in mast cells following
XX activation or in patients with allergic hypersensitivity disease, (I)
XX that encodes proteins (ABN7569-ABN7575) (II) or a protein fragment of
XX (II) if at least 6 amino acids. (II) is useful for identifying binding
XX partners. (I) or (II) is useful for diagnosing or treating a disease
XX state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
XX urticaria or atopic dermatitis or mastocytosis) in a subject which
XX involves determining the level of expression of (I) or (II). A computer
XX system, comprising a database containing information identifying the
XX expression level in a tissue or at least one mast cell of (I), is useful
XX for presenting information to identify the relative expression level of
XX (I). (II) is used as a marker to detect, diagnose or identify an allergic
XX response in a patient. The protein can also serve as a target that
XX modulate gene expression or activity and as an antigen to raise
XX polyclonal or monoclonal antibodies. (II) is useful for identifying
XX agents that modulate expression of the protein or agents, such as
XX agonists or antagonists. The agonists or antagonists are useful for
XX modulating biological activity and function of (II) and thus are useful
XX for alleviating disease conditions such as allergic hypersensitivity,
XX seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.
XX
XX Sequence 1662 BP; 229 A; 551 C; 511 G; 371 T; 0 other;

```

```

Query Match 48.7%; Score 1662; DB 24; Length 1662;
Best Local Similarity 100.0%; Pred No. 7.5e-251;
Matches 1662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 284 ATGTCACAGAGCTGTGGGTGAGCCGCCCTGCTGCGCACCGAAGCCAGCTTTGCTG 343
Db 1 ATGTCACAGAGCTGTGGGTGAGCCGCCCTGCTGCGCACCGAAGCCAGCTTTGCTG 60
Qy 344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTGGCGCAGGCATACCTATGTGCCG 403
Db 61 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTGGCGCAGGCATACCTATGTGCCG 120
Qy 404 CPTCTGCTGTGAAGTGGGGGTAGAGAGAGTTCATGACCATGTGCTGGGCATTGGT 463
Db 121 CPTCTGCTGTGAAGTGGGGGTAGAGAGAGTTCATGACCATGTGCTGGGCATTGGT 180
Qy 464 CCAGTCTGGGCTGTGCTGTGTGCTCCCGCTCTAGGCTCAGCCAGTACCCAGTGGGCTGA 523
Db 181 CCAGTCTGGGCTGTGCTGTGTGCTCCCGCTCTAGGCTCAGCCAGTACCCAGTGGGCTGA 240

```

524	QY	CGCTATGGCGCGCGCGCCCTTCATCTGGGCACCTGTCTCTGGGCGATCTCTGCTGAGCCTC	583
241	Db	CGCTATGGCGCGCGCGCCCTTCATCTGGGCACCTGTCTCTGGGCGATCTCTGCTGAGCCTC	300
584	QY	TTTCTCATCCCAAGGGCGGGTGGCTAGCAGGGCTGCTGTGCCCGGATCCAGGCGCCCTG	643
301	Db	TTTCTCATCCCAAGGGCGGGTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAAGGCGCCCTG	360
644	QY	GAGCTGGGACCTGCTCATCTCGGGCGCTGGGCTGCTGGACTCTGTGGCCAGGTGTGCTTC	703
361	Db	GAGCTGGGACCTGCTCATCTGGGCGTGGGCTGCTGGACTCTGTGGCCAGGTGTGCTTC	420
704	QY	ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCGGGGAOCCGGACCACTGTGCCAGGCC	763
421	Db	ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCGGGGAOCCGGACCACTGTGCCAGGCC	480
764	QY	FACTCTGTCTATGGCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCTGCTGCC	823
481	Db	TACTCTGTCTATGGCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCTGCTGCC	540
824	QY	ATTGACTGGGACACCACTGCCCTGGCCCTTACCTGGGCACCCAGGAGGAGTGCCTCTTT	883
541	Db	ATTGACTGGGACACCACTGCCCTGGCCCTTACCTGGGCACCCAGGAGGAGTGCCTCTTT	600
884	QY	GGCTGCTCAOCCCTCATCTTCTCACTGCTAGCAGCCACACTGCTGGTGGCTGAGGAG	943
601	Db	GGCTGCTCAOCCCTCATCTTCTCACTGCTAGCAGCCACACTGCTGGTGGCTGAGGAG	660
944	QY	GCAGCGCTGGGCCCAACCGAGCCAGGAGAAGGGCTGTGGGCCCTCTCTGTGCGCCAC	1003
661	Db	GCAGCGCTGGGCCCAACCGAGCCAGGAGAAGGGCTGTGGGCCCTCTCTGTGCGCCAC	720
1004	QY	TGCTGTCCATGCCGGCGCGCTTGCTTTCCGGAACCTGGCGCGCCTGCTTCCCGCGGTG	1063
721	Db	TGCTGTCCATGCCGGCGCGCTTGCTTTCCGGAACCTGGCGCGCCTGCTTCCCGCGGTG	780
1064	QY	CACCACTGTGTGCCGCGATGCCCGCACCTCTGCGCGGCTCTTCTGTGCTGAGCTGTGC	1123
781	Db	CACCACTGTGTGCCGCGATGCCCGCACCTCTTCTGTGCTGAGCTGTGC	840
1124	QY	AGCTGGATGGCACTCATGACCTTCACGCTGTGTTTTACACGGATTTCTGTGGCGAGGGGCTG	1183
841	Db	AGCTGGATGGCACTCATGACCTTCACGCTGTGTTTTACACGGATTTCTGTGGCGAGGGGCTG	900
1184	QY	TACAGGGGCTGCCAGAGCTGAGCGGGCACCGAGGCGCGGAGACACTATGATGAAGC	1243
901	Db	TACAGGGGCTGCCAGAGCTGAGCGGGCACCGAGGCGCGGAGACACTATGATGAAGC	960
1244	QY	GTTCCGATGGCAGCCTGGGGGTGTTCTCTCAGTGGCCCATCTCCCTGGTCTCTCTCTG	1303
961	Db	GTTCCGATGGCAGCCTGGGGGTGTTCTCTCAGTGGCCCATCTCCCTGGTCTCTCTCTG	1020
1304	QY	GTCAATGGACCGCTGTGACGGATTCGGCACTCGAGCAGTCTATTTTGGCCAGTGGCA	1363
1021	Db	GTCAATGGACCGCTGTGACGGATTCGGCACTCGAGCAGTCTATTTTGGCCAGTGGCA	1080
1364	QY	GCTTTCCCTGTGGCTGCCGCTGCCATGCTCTCCACAGTGTGGCCGCTGGTGACAGCT	1423
1081	Db	GCTTTCCCTGTGGCTGCCGCTGCCATGCTCTCCACAGTGTGGCCGCTGGTGACAGCT	1140
1424	QY	TCAGCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC	1483
1141	Db	TCAGCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC	1200
1484	QY	TCCTCTACACGGGAGAGCAGGTGTCTTGCCCAATACCGAGGGGACACTGGAGGT	1543
1201	Db	TCCTCTACACGGGAGAGCAGGTGTCTTGCCCAATACCGAGGGGACACTGGAGGT	1260
1544	QY	GCTAGCAGTGGAGACAGCCTGATGACAGCTTCTGCCAGGCCCTAAGCTGGAGCTCCC	1603
1261	Db	GCTAGCAGTGGAGACAGCCTGATGACAGCTTCTGCCAGGCCCTAAGCTGGAGCTCCC	1320
1604	QY	TTCCCTAATGGACACCTGGGTGCTGGAGGCACTGGGCTGCTCCCACTCTCACCCGGGCTC	1663

Db	1321	TTCCCTAAATGGACACAGTGGGTGCTGGAGGCAGTGGCTTGCCTCCACCTCCACCCGCGCTC	1380
Qy	1664	TGCGGGGCTCTCCCTGTGATGTCCTCGTACGTGTGGTGGGTGAGCCACCCAGGCGC	1723
Db	1381	TGCGGGGCTCTCCCTGTGATGTCCTCGTACGTGTGGTGGGTGAGCCACCCAGGCGC	1440
Qy	1724	AGGTGTGTTCCGGCGCGGGCATCTGCCTGGACCTCGCCATCTCGGATAGTGCCTTCCCTG	1783
Db	1441	AGGTGTGTTCCGGCGCGGGCATCTGCCTGGACCTCGCCATCTCGGATAGTGCCTTCCCTG	1500
Qy	1784	CTGTCCAGGTGGCCCATCCCTGTTTATGGGTCTCATTTGTCAGTTCAGCCAGTCTGTC	1843
Db	1501	CTGTCCAGGTGGCCCATCCCTGTTTATGGGTCTCATTTGTCAGTTCAGCCAGTCTGTC	1560
Qy	1844	ACTGCCTATATGCTGTGCCGAGGCGCTGGTCTGGTTCGCAATTTACTTTGCTACACAG	1903
Db	1561	ACTGCCTATATGCTGTGCCGAGGCGCTGGTCTGGTTCGCAATTTACTTTGCTACACAG	1620
Qy	1904	GTAGTATTTGACAGAGCGACTTGGCCAAATACTCAGCGTAG	1945
Db	1621	GTAGTATTTGACAGAGCGACTTGGCCAAATACTCAGCGTAG	1662
RESULT 36			
ABA91283	ID	ABA91283 standard; DNA; 2133 BP.	
XX	AC	ABA91283;	
XX	DT	08-APR-2002 (first entry)	
XX	DE	Thioredoxin-ubiquitin-P501S(aa55-553)His triple gene fusion.	
XX	KW	Thioredoxin; trxA; ubiquitin; P501S; tumour; prostate; antigen;	
XX	KW	cancer; vaccine; therapy; human; gene; ds.	
XX	OS	Chimeric - Homo sapiens.	
XX	OS	Chimeric - Escherichia coli.	
XX	PN	WO200200892-A1.	
XX	PD	03-JAN-2002.	
XX	PF	19-JUN-2001; 2001WO-EP06952.	
XX	PR	26-JUN-2000; 2000GB-0015619.	
XX	PR	30-OCT-2000; 2000GB-0026484.	
XX	PA	(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.	
XX	PI	Cabezon Silva TEV, Delisse AEF;	
XX	DR	WPI; 2002-147888/19.	
XX	P	P-PSDB; AAM50661.	
XX	PT	Novel DNA sequence encoding triple fusion protein comprising ubiquitin	
XX	PT	fused between thioredoxin and polypeptide of interest, useful for	
XX	PT	producing recombinant polypeptide of interest suitable for medicinal	
XX	PT	use	
XX	PS	Example 5; Fig 9a; 87pp; English.	
XX	CC	The present sequence is that of a triple gene fusion comprising,	
XX	CC	from the 5' end, the trxA thioredoxin gene from Escherichia coli,	
XX	CC	the human ubiquitin coding sequence, the prostate antigen P501S	
XX	CC	(amino acids 55-553) coding sequence, and DNA encoding a histidine	
XX	CC	tail. The triple fusion was constructed in plasmid pRIT15063,	
XX	CC	which included the Saccharomyces cerevisiae CUP1 promoter and	
XX	CC	yeast alpha prepro signal sequence. The triple fusion protein	
XX	CC	(see AAM50661) was produced in E. coli GI724 transformants. This	
XX	CC	is an example of the production of triple fusion proteins of the	
XX	CC	invention comprising ubiquitin fused between thioredoxin and a	

CC protein of interest, in this case P501S(aa55-553). A claimed method  
 CC of producing a recombinant protein of interest involves: culturing  
 CC a host cell (preferably *E. coli*) under conditions which allow  
 CC for co-expression of the triple fusion and a ubiquitin-specific  
 CC endoprotease (especially UBP1 from *Saccharomyces cerevisiae*);  
 CC and recovering the recombinant protein directly from the bacterial  
 CC cells after it has been subjected to the action of the  
 CC ubiquitin-specific endoprotease in vivo. In the present case,  
 CC a p501-like protein of 509 amino acids is generated. The  
 CC recombinant protein can used as a vaccine for cancer therapy.  
 XX

SQ Sequence 2133 BP; 384 A; 656 C; 614 G; 479 T; 0 other;

Query Watch 44.1%; Score 1503; DB 24; Length 2133;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-226;  
 Matches 1503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 ATGACCATGGTCTGGCATTTGGTCCAGTCTGGGCGCTGTGTGTCGCCCTCCTAGGC 499  
 DB 604 ATGACCATGGTCTGGCATTTGGTCCAGTCTGGGCGCTGTGTGTCGCCCTCCTAGGC 663

QY 500 TCAGCCAGTACCACTGGCGTGGAGCGCTATGCCCGCCCGCCCTTCACTGGGCACTG 559  
 DB 664 TCAGCCAGTACCACTGGCGTGGAGCGCTATGCCCGCCCGCCCTTCACTGGGCACTG 723

QY 560 TCCTTGGGCATCCTGCTGAGCCTCTTCTCATCCCAAGGGCGGCTGGCTAGCAGGCGTG 619  
 DB 724 TCCTTGGGCATCCTGCTGAGCCTCTTCTCATCCCAAGGGCGGCTGGCTAGCAGGCGTG 783

QY 620 CTGTGCCCGGATCCAGCGCCCTGGAGCTGGCACTGTCTATCCTGGGCGTGGGCTGCTG 679  
 DB 784 CTGTGCCCGGATCCAGCGCCCTGGAGCTGGCACTGTCTATCCTGGGCGTGGGCTGCTG 843

QY 680 GACTTCTGTGGCAGGTGTGCTCACTCCACTGAGGCGCTGTCTGTACCTCTTCCGG 739  
 DB 844 GACTTCTGTGGCAGGTGTGCTCACTCCACTGAGGCGCTGTCTGTACCTCTTCCGG 903

QY 740 GACCCGGACCACTGTGCGCAGGCTACTCTGTCTATGCTTCATGATCATGCTTGGGGC 799  
 DB 904 GACCCGGACCACTGTGCGCAGGCTACTCTGTCTATGCTTCATGATCATGCTTGGGGC 963

QY 800 TGCTTGGGCTACCTCTGCTGCCATTTGACTGGGACACAGTGGCCCTGGCCCTTACCTG 859  
 DB 964 TGCTTGGGCTACCTCTGCTGCCATTTGACTGGGACACAGTGGCCCTGGCCCTTACCTG 1023

QY 860 GGCACCCAGGAGGAGTGCCTTTTGGCGCTGCTACCTCATCTTCCCTACCTGGTAGCA 919  
 DB 1024 GGCACCCAGGAGGAGTGCCTTTTGGCGCTGCTACCTCATCTTCCCTACCTGGTAGCA 1083

QY 920 GCCACACTGTGGTGGCTGAGGAGCGAGCTGGGCGCCACCGAGCCAGAGGAGGCTG 979  
 DB 1084 GCCACACTGTGGTGGCTGAGGAGCGAGCTGGGCGCCACCGAGCCAGAGGAGGCTG 1143

QY 980 TCGGCGCCCTCCTTGTGCGCCCACTGTGTCCATGCGCGGCGCGCTTGGCTTTCCGGAAC 1039  
 DB 1144 TCGGCGCCCTCCTTGTGCGCCCACTGTGTCCATGCGCGGCGCGCTTGGCTTTCCGGAAC 1203

QY 1040 CTGGGCGCCCTGCTTCCCGGCTGCACAGCTGTGCTGGCGCATGCCCGCCAGCTGGCG 1099  
 DB 1204 CTGGGCGCCCTGCTTCCCGGCTGCACAGCTGTGCTGGCGCATGCCCGCCAGCTGGCG 1263

QY 1100 CGGCTCTTGGTGGCTGAGCTGTGAGCTGGATGGCACTCATGACTTTCAGCTGTTTAC 1159  
 DB 1264 CGGCTCTTGGTGGCTGAGCTGTGAGCTGGATGGCACTCATGACTTTCAGCTGTTTAC 1323

QY 1160 ACGATTTCGTGGCGGAGGCGCTATACAGGGCGTGCACAGAGCTGAGCCGGCACCCGAG 1219  
 DB 1324 ACGATTTCGTGGCGGAGGCGCTATACAGGGCGTGCACAGAGCTGAGCCGGCACCCGAG 1383

QY 1220 GCCCGGAGACACTATGATGAAGGCTTCGGATGGGAGCTGGGCTGTTCCTGCAGTGC 1279  
 DB 1384 GCCCGGAGACACTATGATGAAGGCTTCGGATGGGAGCTGGGCTGTTCCTGCAGTGC 1443

QY 1280 GCCATCTCCCTGGTCTTCTCTGTGTCATGACCGGCTGTGTGAGGATTTGGCACTCGA 1339  
 DB 1444 GCCATCTCCCTGGTCTTCTCTGTGTCATGACCGGCTGTGTGAGGATTTGGCACTCGA 1503

QY 1340 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGTCCCGGTCACATGCTGTGTC 1399  
 DB 1504 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGTCCCGGTCACATGCTGTGTC 1563

QY 1400 CACAGTGTGGCGTGGTGTGACAGCTTCAGCGCGCTTCACCGGGTTCACCTTCTCAGCCCTG 1459  
 DB 1564 CACAGTGTGGCGTGGTGTGACAGCTTCAGCGCGCTTCACCGGGTTCACCTTCTCAGCCCTG 1623

QY 1460 CAGATCTGCCCTTACACACTGGGCTCCCTCTTACACCGGGAGAGAGTGTTCCTGCCC 1519  
 DB 1624 CAGATCTGCCCTTACACACTGGGCTCCCTCTTACACCGGGAGAGAGTGTTCCTGCCC 1683

QY 1520 AATACCGGAGGACACTGGAGGTGTAGCAGTGTGAGGACAGCTGATGACCACTTCTCTG 1579  
 DB 1684 AATACCGGAGGACACTGGAGGTGTAGCAGTGTGAGGACAGCTGATGACCACTTCTCTG 1743

QY 1580 CCAGGCGCTAAGCCTGGAGCTCCCTTCCCTTAATGGACAGTGGTGTCTGGAGGCACTGGC 1639  
 DB 1744 CCAGGCGCTAAGCCTGGAGCTCCCTTCCCTTAATGGACAGTGGTGTCTGGAGGCACTGGC 1803

QY 1640 CTGCTCCAGCTCCACCGCGCTCTCGGGGCTCTGCGGCGCTCTGCTGTGATGTCTCCGTACGTG 1699  
 DB 1804 CTGCTCCAGCTCCACCGCGCTCTCGGGGCTCTGCGGCGCTCTGCTGTGATGTCTCCGTACGTG 1863

QY 1700 GTGGTGGGTGAGCCACCGAGGCGAGGTGGTTCGGGCGCGGCGCATCTGCTGGACCTC 1759  
 DB 1864 GTGGTGGGTGAGCCACCGAGGCGAGGTGGTTCGGGCGCGGCGCATCTGCTGGACCTC 1923

QY 1760 GCCATCTGGATAGTCCCTTCCCTGTGTGCCAGGTGGCGCCCATCCCTGTTTATGGGCTCC 1819  
 DB 1924 GCCATCTGGATAGTCCCTTCCCTGTGTGCCAGGTGGCGCCCATCCCTGTTTATGGGCTCC 1983

QY 1820 ATTGTCCAGCTCAGCCAGTCTGTCACTGCTATATGCTGTGCGCCAGGCGCTGGGTCTG 1879  
 DB 1984 ATTGTCCAGCTCAGCCAGTCTGTCACTGCTATATGCTGTGCGCCAGGCGCTGGGTCTG 2043

QY 1880 GTGCCCATTTACTTTCTCTACAGAGTATTTGACAAGAGGCACTTGGCCAAATCTCA 1939  
 DB 2044 GTGCCCATTTACTTTCTCTACAGAGTATTTGACAAGAGGCACTTGGCCAAATCTCA 2103

QY 1940 GCG 1942  
 DB 2104 GCG 2106

## RESULT 37

ABA91284

ID ABA91284 standard; DNA; 1593 BP.

XX ABA91284;

AC AC

XX DT

XX 08-APR-2002 (first entry)

DE Thioresoxin-ubiquitin-p501S(aal-320)His triple gene fusion.

XX Thioresoxin; trxA; ubiquitin; p501S; tumour; prostate; antigen;

KW cancer; vaccine; therapy; human; gene; ds.

XX Chimeric - Homo sapiens.

OS Chimeric - Escherichia coli.

XX

FH Key

FT 1..1593

FT CDS

FT /tag= a

FT /transl\_except= (pos:607..609,aa:Xaa)

FT Figure 12b (AA50662), which has a Val

FT residue at amino acid position 208 not

FT not encoded by the present sequence"

FT









The present invention describes polynucleotide sequences (I) which encode prostate-specific proteins (II). (I) and (II) have cytostatic activity, and can be used in vaccine production and gene therapy. (I), (II), antibodies to (II), fusion proteins comprising (II), and isolated T cells prepared using (I) or (II) are used treat cancer in a patient. (I) and the antibodies are also used in the detection of cancer in a patient. The cancer that is diagnosed or treated is particularly prostate cancer. (I) and (II) can be used in vaccines. The antibodies or (I) can be used for monitoring the progression of cancer in a patient. (I) and (II) can also be used to improve diagnostic and therapeutic methods for prostate cancer. They can indicate the level of metastasis as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to AAH01318 represent polynucleotide and amino acid sequences used in the exemplification of the present invention.

Sequence 1203 BP; 175 A; 415 C; 368 G; 245 T; 0 other;

Query Match	Score	DB	Length
23.38	794.6	DB	1203

Query Match 23.3%; Score 794.0; DB 22; Length 1203;  
Best Local Similarity 99.58; Pred. No. 2.2e-115;

BEST LOCAL SIMILARITY 99.3%; PRED: NO. 2.2E-115;  
 Matches 797; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

379	QY	GGCCGACGGCATCACCTATATGTGCCGCTCTGCTGCTGGAGTGGGGGTAGAGGAAGTT	438
402	Db		
402	Db	GGCGAAATTCACTACCTATATGTGCCGCTCTGCTGCTGGAAGTGGGGTAGAGGAAGTT	461
439	QY		
439	QY	CATGACCATGTGCTGGGCATTGGTCCAGTGCTGGGCGTGCTGTGTCGCCGCTCCTAGG	498
462	Db		
462	Db	CATGACCATGTGCTGGGCATTGGTCCAGTGCTGGGCGCTGGTCTGTGTCGCCGCTCCTAGG	521
499	QY		
499	QY	CTCAGCCAGTGACCACTGGGCTGGAGCGTATGSCCGCGCCGCGGCTTCATCTGGGCAC	558
522	Db		
522	Db	CTCAGCCAGTGACCACTGGGCTGGAGCGTATGSCCGCGCGCGGCTTCATCTGGGCAC	581
559	QY		
559	QY	GTCTTGGGCATCTGCTGAGCGCTTTTCTCATCCCAAGGGCGGCTGGCTAGCAGGGCT	618
582	Db		
582	Db	GTCTTGGGCATCTGCTGAGCGCTTTTCTCATCCCAAGGGCGGCTGGCTAGCAGGGCT	641
619	QY		
619	QY	GCTGTGCCGGATCCCAAGGCCCTGGAGCTGGGACGTGCTCATCTGGGCGTGGGGCTGCT	678
642	Db		
642	Db	GCTGTGCCGGATCCCAAGGCCCTGGAGCTGGGACGTGCTCATCTGGGCGTGGGGCTGCT	701
679	QY		
679	QY	GGACTTCTGTGGCCAGGTGCTTCACTCCACTGGAGGCCCTGCTCTGTGACCTCTTCCG	738
702	Db		
702	Db	GGACTTCTGTGGCCAGGTGCTTCACTCCACTGGAGGCCCTGCTCTGTGACCTCTTCCG	761
739	QY		
739	QY	GGACCCGACCACTGTGCGCAGGCGCTACTGTCTATATGCCCTCATGATCACTTTGGGG	798
762	Db		
762	Db	GGACCCGACCACTGTGCGCAGGCGCTACTGTCTATATGCCCTCATGATCACTTTGGGG	821
799	QY		
799	QY	CTGCTGTGGCTTACCTCTGCTGCCATTGACTGGGACACAGTGCCTGGCCCTACCT	858
822	Db		
822	Db	CTGCTGTGGCTTACCTCTGCTGCCATTGACTGGGACACAGTGCCTGGCCCTACCT	881
859	QY		
859	QY	GGGCACCCAGGAGAGTGCCTTTTGGCCTGTCACTTCCCTACCTCGCTAGC	918
882	Db		
882	Db	GGGCACCCAGGAGAGTGCCTTTTGGCCTGTCACTTCCCTACCTCGCTAGC	941
919	QY		
919	QY	AGCCACACTGCTGTGGCTGAGAGAGCGGGCTGGGCCCCACCGACGACGAGAAGGCT	978
942	Db		
942	Db	AGCCACACTGCTGTGGCTGAGAGAGCGGGCTGGGCCCCACCGACGACGAGAAGGCT	1001
979	QY		
979	QY	GTGGGCCCTTCTTGTGCGCCCACTGCTGTCCATGCGGGGCGCGCTTTCGGGAA	1038
1002	Db		
1002	Db	GTGGGCCCTTCTTGTGCGCCCACTGCTGTCCATGCGGGGCGCGCTTTCGGGAA	1061
1039	QY		
1039	QY	CCTGGGGCCCTGCTTCCCGGGGTGACCAAGCTGTGCTGCCGATGCCCGCACCTCGC	1098
1062	Db		
1062	Db	CCTGGGGCCCTGCTTCCCGGGGTGACCAAGCTGTGCTGCCGATGCCCGCACCTCGC	1121
1099	QY		
1099	QY	CGGCTCTTGTGGCTGAGCTGTGACGTGGATGGCACTCATGACCTTCACGCTGTTTA	1158
1122	Db		
1122	Db	CGGCTCTTGTGGCTGAGCTGTGACGTGGATGGCACTCATGACCTTCACGCTGTTTA	1181

XX  
PT  
PT  
PT  
PS  
PS  
XX  
XX  
CC  
CC  
CC  
CC  
XX  
SQ

New prostate-specific polynucleotides for diagnosing and treating diseases, in particular prostate cancer, and as markers for the progression of cancer

Example 17; SEQ ID NO 851; 87pp; English.

The present invention provides prostate-specific coding sequences and their encoded proteins. These can be used in the diagnosis and treatment of cancers, particularly prostate cancer. The present sequence is a cDNA described in the invention.

Sequence 1203 BP; 175 A; 415 C; 368 G; 245 T; 0 other;

Query Match 23.3%; Score 794.6; DB 24; Length 1203;  
Best Local Similarity 99.5%; Pred. No. 2.2e-115;  
Matches 797; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

379 GGCAGGAGGATACCTATGTCGCCCTCTGCTGGAAGTGGGGGTAGAGAGATT 438  
|||||  
402 GCGCGAATTCATCACTATGTCGCCCTCTGCTGGAAGTGGGGGTAGAGAGATT 461  
439 CATGACCATGCTGGGATGTCAGTGTGGGCTGTCTGTCTGCTCCGCTCTTAGG 498  
|||||  
462 CATGACCATGCTGGGATGTCAGTGTGGGCTGTCTGTCTGCTCCGCTCTTAGG 521  
499 CTCAGCAGTACCACTGGCGTGGAGCTATGCGCGCGCGCCCTTCATCTGGCACT 558  
|||||  
522 CTCAGCAGTACCACTGGCGTGGAGCTATGCGCGCGCGCCCTTCATCTGGCACT 581  
559 GTCTTTGGCATCTGCTGAGGCTCTTTCTCATCCCAAGGCGCGCTGCTAGCAGGCT 618  
|||||  
582 GTCTTTGGCATCTGCTGAGGCTCTTTCTCATCCCAAGGCGCGCTGCTAGCAGGCT 641  
619 GCTGTCCCGGATCCAGGCGCTGAGCTGCTGCTCATCTGCTGCTGCTGCTGCTGCT 678  
|||||  
642 GCTGTCCCGGATCCAGGCGCTGAGCTGCTGCTCATCTGCTGCTGCTGCTGCTGCT 701  
679 GAGCTTCTGTGCGAGTGTCTTCACTCCAGGCGCGCTGCTGCTGCTGCTGCTGCTGCT 738  
|||||  
702 GAGCTTCTGTGCGAGTGTCTTCACTCCAGGCGCGCTGCTGCTGCTGCTGCTGCTGCT 761  
739 GAGCCCGGACCACTGCTGCGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798  
|||||  
762 GAGCCCGGACCACTGCTGCGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 821  
799 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858  
822 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 881  
859 GGCACCCAGGAGGAGTGTCTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 918  
|||||  
882 GGCACCCAGGAGGAGTGTCTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 941  
919 AGCCACACCTGCTGGTGGCTGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978  
942 AGCCACACCTGCTGGTGGCTGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001  
979 GTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038  
1002 GTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1061  
1039 CTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098  
1062 CTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121  
1099 CCGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1158  
1122 CCGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181  
1159 CACGGATTTCTGCTGGGCGAGG 1179  
1182 CACGGATTTCTGCTGGGCGAGTG 1202

RESULT 42  
AAV61144  
ID AAV61144 standard; cDNA; 789 BP.  
XX  
AC AAV61144;  
XX  
DT 06-JAN-1999 (first entry)  
XX  
DE 3' cDNA sequence of prostate tumour clone L1-12.  
XX  
KW Prostate; cancer; tumour; vaccine; immunogen; clone; ss.  
XX  
OS Homo sapiens.  
XX  
PN W09837093-A2.  
XX  
PD 27-AUG-1998.  
XX  
PF 25-FEB-1998; 98WO-US03492.  
XX  
PR 09-FEB-1998; 98US-0020956.  
XX  
PR 25-FEB-1997; 97US-0806099.  
XX  
PR 01-AUG-1997; 97US-0904804.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Dillion DC, Xu J;  
XX  
XX WPI; 1998-609886/51.  
XX  
PT Polypeptides comprising immunogenic portions of prostate proteins used in a vaccine for the treatment of prostate cancer  
XX  
PS Claim 3; Page 38-39; 130pp; English.  
XX  
CC The present sequence is a new DNA which encodes an immunogenic portion of a prostate tumour protein. The encoded immunogen, or the DNA itself, can be used as a vaccine for the treatment of prostate cancer. The DNA was identified by analysis of a subtracted cDNA library obtained by subtracting a prostate tumour cDNA expression library with a normal tissue cDNA library.  
XX  
SQ Sequence 789 BP; 122 A; 250 C; 221 G; 182 T; 14 other;

Query Match 19.7%; Score 673.4; DB 19; Length 789;  
Best Local Similarity 94.2%; Pred. No. 1.8e-96;  
Matches 745; Conservative 0; Mismatches 40; Indels 6; Gaps 5;

1341 CAGTCTATTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGGCGGTGCCACATGCTGCTGCC 1400  
|||||  
1 CAGTCTATTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGGCGGTGCCACATGCTGCTGCC 60  
1401 ACAGTGTGGCCGTGGTGTACAGCTTTACGCCCTTACCGGGTTACCTTTCAGCCCTGC 1450  
|||||  
61 ACAGTGTGGCCGTGGTGTACAGCTTTACGCCCTTACCGGGTTACCTTTCAGCCCTGC 120  
1461 AGATCTCCCTACACACTGCTGCTTACACCGGGGAGAGAGAGTGTTCCTGCCCA 1520  
|||||  
121 AGATCTCCCTACACACTGCTGCTTACACCGGGGAGAGAGAGTGTTCCTGCCCA 180  
1521 AATACCGAGGAGAGAGTGTGCTAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1580  
|||||  
181 AATACCGAGGAGAGAGTGTGCTAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
1581 CAGGCCCTTAAGCTGGAGCTCCCTTCCCTTAATGACACAGTGGGTGCTGGAGCAGTGGCC 1640  
|||||  
241 CAGGCCCTTAAGCTGGAGCTCCCTTCCCTTAATGACACAGTGGGTGCTGGAGCAGTGGCC 300  
1641 TGCTCCCACTCCACCCCGCTCTGCGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700  
|||||  
301 TGCTCCCACTCCACCCCGCTCTGCGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 13:11:18 ; Search time 8483 Seconds  
(without alignments)  
11698.758 Million cell updates/sec

Title: US-09-759-143-110

Perfect score: 3410  
Sequence: 1 ggggaaccagctgcacgcgc.....aaaaaaaaaaaaaaaa 3410.

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_ro:\*  
11: gb\_ro:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vt:\*  
30: em\_hgt\_hum:\*  
31: em\_hgt\_inv:\*  
32: em\_hgt\_other:\*  
33: em\_hgt\_mus:\*  
34: em\_hgt\_pln:\*  
35: em\_hgt\_rod:\*  
36: em\_hgt\_mam:\*  
37: em\_hgt\_vrt:\*  
38: em\_sy:\*  
39: em\_hgt\_hum:\*  
40: em\_hgt\_mus:\*  
41: em\_hgt\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3409.6	100.0	3410	6	AX106329	Sequence
2	3409.6	100.0	3410	6	AX140620	Sequence
3	3409.6	100.0	3410	6	AX200480	Sequence
4	3409.6	100.0	3410	6	AX267136	Sequence
5	3409.6	100.0	3410	6	AX267136	Sequence
6	3409.6	100.0	3410	6	AX267136	Sequence
7	3292.4	96.6	3320	9	AY033593	Homo sapi
8	2900.2	85.0	3514	9	AB060851	Macaca fa
9	2585.4	75.8	4034	6	AX200995	Sequence
10	2585.4	75.8	4034	6	AX267730	Sequence
11	2545.6	74.7	2917	9	AB062977	Macaca fa
12	2196.4	64.4	2904	6	AX200994	Sequence
13	2196.4	64.4	2904	6	AX267729	Sequence
14	2142.8	62.8	4894	6	AX200993	Sequence
15	2142.8	62.8	4894	6	AX267728	Sequence
16	2136.4	62.7	2152	6	AR112295	Sequence
17	2114.8	62.0	2143	6	AR112294	Sequence
18	2103.8	61.7	2477	9	HS080424	Homo sapi
19	1815.8	53.2	6976	6	AX200996	Sequence
20	1815.8	53.2	6976	6	AX267731	Sequence
21	1809.8	53.1	157988	2	AC096533	Homo sapi
22	1583.4	46.4	3354	10	BC031381	Mus muscu
23	1503	44.1	2133	6	AX343857	Sequence
24	1246.2	36.5	2611	10	BC034084	Mus muscu
25	961.2	28.2	1593	6	AX343860	Sequence
26	794.6	23.3	1203	6	AX201078	Sequence
27	794.6	23.3	1203	6	AX267877	Sequence
28	673.4	19.7	789	6	AX106229	Sequence
29	673.4	19.7	789	6	AX140520	Sequence
30	673.4	19.7	789	6	AX200380	Sequence
31	673.4	19.7	789	6	AX267036	Sequence
32	633.2	18.6	198037	2	AC126523	Rattus no
33	623.6	18.3	178567	2	AC107837	Mus muscu
34	603	17.7	772	6	AX106230	Sequence
35	603	17.7	772	6	AX140521	Sequence
36	603	17.7	772	6	AX200381	Sequence
37	603	17.7	772	6	AX267037	Sequence
38	442.6	13.0	1593	10	BC024519	Mus muscu
39	361.8	10.6	198037	2	AC126523	Rattus no
40	319.8	9.4	342	6	AR112287	Sequence
41	291	8.5	294	6	AR112291	Sequence
42	287	8.4	288	6	AR112289	Sequence
43	270	7.9	272	6	AR112290	Sequence
44	265	7.8	265	6	AR112288	Sequence
45	261.4	7.7	406	11	HSPE54C06	AL033941 H.sapiens

# ALIGNMENTS

RESULT 1  
AX106329  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 3410)  
XU,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.  
Compositions and methods for therapy and diagnosis of prostate cancer

AX106329  
Sequence 110 from Patent WO0125272.  
AX106329  
GI:13922014

3410 bp  
DNA  
linear  
PAT 30-APR-2001

Pred. No. is the number of results predicted by chance to have a

JOURNAL	Patent: WO 0125272-A 110 12-APR-2001;
FEATURES	CORIXA CORPORATION (US)
source	Location/Qualifiers
	1. .3410
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	667 a 1014 c 945 g 783 t 1 others
ORIGIN	
Query Match	100.0%; Score 3409.6; DB 6; Length 3410;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3410; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GGGAAACAGCCTGCACGGCTGGCTCCGGGTGACAGCCGCGGCTCGGCGAGATCTGA 60
Db	1 GGGAAACAGCCTGCACGGCTGGCTCCGGGTGACAGCCGCGGCTCGGCGAGATCTGA 60
QY	61 GTGATGAGACCTGTCCTCCACTAGCTGAGTCCCGCCACAGCAGAGGTGTTGAGCATGGGCTGAG 120
Db	61 GTGATGAGACCTGTCCTCCACTAGCTGAGTCCCGCCACAGCAGAGGTGTTGAGCATGGGCTGAG 120
QY	121 AAGCTGGACCGCCACCAAGGCTGGCAGAAATGGCGGCTTGGCTGATTCCTAGGCAGTT 180
Db	121 AAGCTGGACCGCCACCAAGGCTGGCAGAAATGGCGGCTTGGCTGATTCCTAGGCAGTT 180
QY	181 GCGGCGACAGAGAGAGAGCCGACGCTTCTGGAGCAGAGCCGAGCAGAGCAGTTCG 240
Db	181 GCGGCGACAGAGAGAGAGCCGACGCTTCTGGAGCAGAGCCGAGCAGAGCAGTTCG 240
QY	241 GAGTGCCTGAACGCGCCCTGAGCCCTACCGGCTGGCCGCTATGTTCCAGAGGCTGTG 300
Db	241 GAGTGCCTGAACGCGCCCTGAGCCCTACCGGCTGGCCGCTATGTTCCAGAGGCTGTG 300
QY	301 GGTGAGCGCCTGCTGGGACACCGAAAGCCAGCTTCTGCTGATCAACCTGCTAACTT 360
Db	301 GGTGAGCGCCTGCTGGGACACCGAAAGCCAGCTTCTGCTGATCAACCTGCTAACTT 360
QY	361 TGGCTGAGGTGTTGGCCGAGCAGCATCTATGTCGCCCTCTGCTGCTGAGT 420
Db	361 TGGCTGAGGTGTTGGCCGAGCAGCATCTATGTCGCCCTCTGCTGCTGAGT 420
QY	421 GGGGTAGAGAGAGTTCATGACCATGGTGGGCTTGGTCCAGTGTGGGCTGGT 480
Db	421 GGGGTAGAGAGAGTTCATGACCATGGTGGGCTTGGTCCAGTGTGGGCTGGT 480
QY	481 CTGTGTCGGCTCTAGCTGACCATGACCATGGCTGGGCTGAGGCTATGGCGCGCGCG 540
Db	481 CTGTGTCGGCTCTAGCTGACCATGACCATGGCTGGGCTGAGGCTATGGCGCGCGCG 540
QY	541 GCCCTTCATCTGGGCACTGTCTTGGGCACTCTGCTGAGCCTCTTCTCATCCCAAGGGC 600
Db	541 GCCCTTCATCTGGGCACTGTCTTGGGCACTCTGCTGAGCCTCTTCTCATCCCAAGGGC 600
QY	601 CGGCTGCTAGCAGGGCTGTGTGCCCGGATCCCGAGCCCTGGAGTGGCACTGCTCAT 660
Db	601 CGGCTGCTAGCAGGGCTGTGTGCCCGGATCCCGAGCCCTGGAGTGGCACTGCTCAT 660
QY	661 CCTGGGCTGGGCTGTGACATCTCTGGCAGGTGTGTTCACTCCACTGGAGGCTT 720
Db	661 CCTGGGCTGGGCTGTGACATCTCTGGCAGGTGTGTTCACTCCACTGGAGGCTT 720
QY	721 GCTCTGACCTCTTCGGGACCGGACCACTGTGCGCAGGCTACTCTGCTATGCTT 780
Db	721 GCTCTGACCTCTTCGGGACCGGACCACTGTGCGCAGGCTACTCTGCTATGCTT 780
QY	781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCT 840
Db	781 CATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCT 840
QY	841 TGCCCTGGCCCCCTTACCTGGGACCCAGAGGAGTGTGCTTCTGGGCTGTCTACCTCAT 900
Db	841 TGCCCTGGCCCCCTTACCTGGGACCCAGAGGAGTGTGCTTCTGGGCTGTCTACCTCAT 900
QY	901 CTTCTCTACCTCGGTAGCAGCCACACTGCTGTGTGCTGAGGAGGAGCGCTGGGCCCCAC 960
Db	901 CTTCTCTACCTCGGTAGCAGCCACACTGCTGTGTGCTGAGGAGGAGCGCTGGGCCCCAC 960
QY	961 CGAGCCAGCAGAGGCTCTCGGCCCCCTCTCTGCGCCCCCTCTCTGCTGCTGCTGCTGCT 1020
Db	961 CGAGCCAGCAGAGGCTCTCGGCCCCCTCTCTGCGCCCCCTCTCTGCTGCTGCTGCTGCT 1020
QY	1021 CGGCTTGGCTTTCCGGAACCTGGGGCCCTGCTTCCCGGCTGCACAGCTGTGCTGCTGCT 1080
Db	1021 CGGCTTGGCTTTCCGGAACCTGGGGCCCTGCTTCCCGGCTGCACAGCTGTGCTGCTGCT 1080
QY	1081 CATGCCCCGACCCCTGCGCGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db	1081 CATGCCCCGACCCCTGCGCGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY	1141 GACCTTCACGCTGTTTACACGATTTCTGCGGCGAGGGGCTGTACAGGGGCTGCCAG 1200
Db	1141 GACCTTCACGCTGTTTACACGATTTCTGCGGCGAGGGGCTGTACAGGGGCTGCCAG 1200
QY	1201 AGCTAGCCGGGACCCGAGGCGCCGAGCATTATGATGAAGGCTTTCGATGGGAGGCT 1260
Db	1201 AGCTAGCCGGGACCCGAGGCGCCGAGCATTATGATGAAGGCTTTCGATGGGAGGCT 1260
QY	1261 GGGCTGTTCTGAGTGGGCTATCTCCCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCT 1320
Db	1261 GGGCTGTTCTGAGTGGGCTATCTCCCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCT 1320
QY	1321 GCAGCGATTTCGACCTCGAGCAGTCTATTGCGCAGTGTGGCAGCTTTCCCTGCTGCTG 1380
Db	1321 GCAGCGATTTCGACCTCGAGCAGTCTATTGCGCAGTGTGGCAGCTTTCCCTGCTGCTG 1380
QY	1381 CGGTGCCACATGCTTCCACAGTGTGGCCTGTGTGACAGCTTCAGCGCGCCCTCACCG 1440
Db	1381 CGGTGCCACATGCTTCCACAGTGTGGCCTGTGTGACAGCTTCAGCGCGCCCTCACCG 1440
QY	1441 GTTCACTTCTCAGCCCTGACATGCTGCTTACACTGAGCTGCTGCTGCTGCTGCTGCT 1500
Db	1441 GTTCACTTCTCAGCCCTGACATGCTGCTTACACTGAGCTGCTGCTGCTGCTGCTGCT 1500
QY	1501 GAAGCAGGTGTTCTGCCCCAAATACGAGGGGACACTGAGGCTGTGAGCAGTGTGAGCAG 1560
Db	1501 GAAGCAGGTGTTCTGCCCCAAATACGAGGGGACACTGAGGCTGTGAGCAGTGTGAGCAG 1560
QY	1561 CCTGATGACCATGCTTCTGCGCAGGCTTAAAGCTGAGCTTCCCTTCCCTTAAATGAGCAG 1620
Db	1561 CCTGATGACCATGCTTCTGCGCAGGCTTAAAGCTGAGCTTCCCTTCCCTTAAATGAGCAG 1620
QY	1621 GGTGCTGAGCAGTGGCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1680
Db	1621 GGTGCTGAGCAGTGGCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1680
QY	1681 TGATGCTTCCCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740
Db	1681 TGATGCTTCCCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1740
QY	1741 GGGCATCTGCTGAGCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db	1741 GGGCATCTGCTGAGCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY	1801 ATCCCTGTTTATGGGCTCCATTTGCTAGCTAGCAGTGTGCTGCTGCTGCTGCTGCTG 1860
Db	1801 ATCCCTGTTTATGGGCTCCATTTGCTAGCTAGCAGTGTGCTGCTGCTGCTGCTGCTG 1860
QY	1861 TGCCGAGGCTTGGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920
Db	1861 TGCCGAGGCTTGGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1920
QY	1921 CGACTTGGCCAAATACTCAGCGTAGAAAACCTTCCAGCAGATTTGGGGTGGAGGCTGCT 1980
Db	1921 CGACTTGGCCAAATACTCAGCGTAGAAAACCTTCCAGCAGATTTGGGGTGGAGGCTGCT 1980
QY	1981 CACTGGTCCAGCTCCCGGCTCCTGTTAGCCCCCATGGGCTGGCGGCTGCCGCCAGT 2040



QY	301	GGT	GAGCCGCGCTGCTGCGGCACCGGAAGCCAGCTCTTGCTGCTCAACTGTAACCTT	360
DB	301	GGT	GAGCCGCGCTGCTGCGGCACCGGAAGCCAGCTCTTGCTGCTCAACTGTAACCTT	360
QY	361	TGGCCT	GGAGGTGTGTTTGGCCGAGGCATCACTATGTGCCGCTCTGCTGCTGGAAGT	420
DB	361	TGGCCT	GGAGGTGTGTTTGGCCGAGGCATCACTATGTGCCGCTCTGCTGCTGGAAGT	420
QY	421	GGGGT	TAGAGGAGAAGTTTATGACATGGTGTGGGCATTTGGTCCAGTGTGGGCTTGGT	480
DB	421	GGGGT	TAGAGGAGAAGTTTATGACATGGTGTGGGCATTTGGTCCAGTGTGGGCTTGGT	480
QY	481	CTGTGT	CCCCGCTCTTAGGCTCAGCCAGTGCACCACTGGCGTGGACGTATATGGCCGCGCGG	540
DB	481	CTGTGT	CCCCGCTCTTAGGCTCAGCCAGTGCACCACTGGCGTGGACGTATATGGCCGCGCGG	540
QY	541	GCCCTT	CATCTGGGCACGTGCTTGGGCATCTGCTGAGCCCTTCTTCATCCCAAGGCG	600
DB	541	GCCCTT	CATCTGGGCACGTGCTTGGGCATCTGCTGAGCCCTTCTTCATCCCAAGGCG	600
QY	601	CGGCT	TGGCTAGCAGGGCTGTGTGCCCGGATCCCAAGCCCTTGGAGCTGSCACTGTCTCAT	660
DB	601	CGGCT	TGGCTAGCAGGGCTGTGTGCCCGGATCCCAAGCCCTTGGAGCTGSCACTGTCTCAT	660
QY	661	CCTGGGCGT	GGGGCTGTGACTTCTGTGGCCAGGTGTCTCACTCCACTGAGAGGCCCT	720
DB	661	CCTGGGCGT	GGGGCTGTGACTTCTGTGGCCAGGTGTCTCACTCCACTGAGAGGCCCT	720
QY	721	GCTCTCT	GCACCTCTTCCGGGACCCGGACCACTGTGCGCAGGCCCTACTCTCTCATGCTT	780
DB	721	GCTCTCT	GCACCTCTTCCGGGACCCGGACCACTGTGCGCAGGCCCTACTCTCTCATGCTT	780
QY	781	CATGAT	CAGTCTTGGGGCTGCTTGGGCTACCTCTGCTTCCCTGCCATTTGACTGGGACACCAG	840
DB	781	CATGAT	CAGTCTTGGGGCTGCTTGGGCTACCTCTGCTTCCCTGCCATTTGACTGGGACACCAG	840
QY	841	TGCCCT	TGGCCCTTACCTTGCGCACCCAGGAGGTGCTCTTTTGGGCTGCTCACTCCCTCAT	900
DB	841	TGCCCT	TGGCCCTTACCTTGCGCACCCAGGAGGTGCTCTTTTGGGCTGCTCACTCCCTCAT	900
QY	901	CTTCCCT	CACCTGCGTAGCAGCCACACTGCTTGGTGGCTGAGGAGCAGCGTGGGCCCCAC	960
DB	901	CTTCCCT	CACCTGCGTAGCAGCCACACTGCTTGGTGGCTGAGGAGCAGCGTGGGCCCCAC	960
QY	961	CGAGCC	AGCAGAGGGCTGTGCGGCCCTCCTTGTTCGCCCCACTGCTGTCCATCGCGGGC	1020
DB	961	CGAGCC	AGCAGAGGGCTGTGCGGCCCTCCTTGTTCGCCCCACTGCTGTCCATCGCGGGC	1020
QY	1021	CCGCTT	TGGCTTCCGGAACCTGGGCGCCCTGCTTCCCGGCTGACAGTGTGCTGTCCG	1080
DB	1021	CCGCTT	TGGCTTCCGGAACCTGGGCGCCCTGCTTCCCGGCTGACAGTGTGCTGTCCG	1080
QY	1081	CATGCC	CCGCACCTTGCGCGGCTCTTCTGTGGTGGCTGAGCTGTGACGTGGATGGCACTCAT	1140
DB	1081	CATGCC	CCGCACCTTGCGCGGCTCTTCTGTGGTGGCTGAGCTGTGACGTGGATGGCACTCAT	1140
QY	1141	GACCTT	CACGCTGTTTTACACGGATTTCTGTGGCGAGGGGCTGTACACAGGCGTGGCCAG	1200
DB	1141	GACCTT	CACGCTGTTTTACACGGATTTCTGTGGCGAGGGGCTGTACACAGGCGTGGCCAG	1200
QY	1201	AGCTGAGC	CGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCCGATGGGCGAGCCT	1260
DB	1201	AGCTGAGC	CGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCCGATGGGCGAGCCT	1260
QY	1261	GGGCGT	TGCTTGCAGTGGCCCATCTCCCTGGTCTCTCTCTGGTCTATGGACCGGTGGT	1320
DB	1261	GGGCGT	TGCTTGCAGTGGCCCATCTCCCTGGTCTCTCTCTGGTCTATGGACCGGTGGT	1320
QY	1321	GCAGCGAT	TGGGACTCGACAGTCTATTTGGCCAGTGTGGGAGCTTCCCTGTGCTGC	1380
DB	1321	GCAGCGAT	TGGGACTCGACAGTCTATTTGGCCAGTGTGGGAGCTTCCCTGTGCTGC	1380
QY	1381	CGGTG	CCACATGCTGTCCACAGTGTGGCGGCTGGTGACAGCTTCAGCCGCCCTCACCGG	1440

Db	1381		CGGTGCCACATGCGCTGCCACAGTGTGGCGTGTGTCACAGCTTACGCCGCTCACCGG	1440
Qy	1441		GTTTACACTTTCACAGCCCTGCAGATCCCTGCCCTACACACTGGCCCTCCCTCTACACACGGGA	1500
Db	1441		GTTTACACTTTCACAGCCCTGCAGATCCCTGCCCTACACACTGGCCCTCCCTCTACACACGGGA	1500
Qy	1501		GAAGCAGTGTTCCTGTCGCCAAATACGAGAGGGGACACTGGAGGTGCTAGCAGTCAGGACAG	1560
Db	1501		GAAGCAGTGTTCCTGTCGCCAAATACGAGAGGGGACACTGGAGGTGCTAGCAGTCAGGACAG	1560
Qy	1561		CCTGATGACCAGCTTCTTGCAGGCCCTAAGCCCTGGAGCTCCCTTCCCTAATGGACACGT	1620
Db	1561		CCTGATGACCAGCTTCTTGCAGGCCCTAAGCCCTGGAGCTCCCTTCCCTAATGGACACGT	1620
Qy	1621		GGTGTCTGGAGGCACTGGCCCTGCTCCACCTCCACCCGCGCTCTGCGGGGCCCTTGCCCTG	1680
Db	1621		GGTGTCTGGAGGCACTGGCCCTGCTCCACCTCCACCCGCGCTCTGCGGGGCCCTTGCCCTG	1680
Qy	1681		TGATGTCTCCGTACGTGTGTGTGGTGTGAGCCACCGAGGCCAGGGTGTGTCGGGGCCG	1740
Db	1681		TGATGTCTCCGTACGTGTGTGTGGTGTGAGCCACCGAGGCCAGGGTGTGTCGGGGCCG	1740
Qy	1741		GGGATCTGCCGTGACCTCGCCATCTCTGATGTGCTTACACAGTACTATTTGACAAGAG	1800
Db	1741		GGGATCTGCCGTGACCTCGCCATCTCTGATGTGCTTACACAGTACTATTTGACAAGAG	1800
Qy	1801		ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCCTATATGTGTCT	1860
Db	1801		ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCCTATATGTGTCT	1860
Qy	1861		TGCCGCAAGCCCTGGGTCTGGTCGCCATTTACTTTGCTTACACAGTACTATTTGACAAGAG	1920
Db	1861		TGCCGCAAGCCCTGGGTCTGGTCGCCATTTACTTTGCTTACACAGTACTATTTGACAAGAG	1920
Qy	1921		CGACTTGGCCAAATACTACAGCGTAGAAAACTTCCAGCACATTTGGGTGGAGGCCCTGCCT	1980
Db	1921		CGACTTGGCCAAATACTACAGCGTAGAAAACTTCCAGCACATTTGGGTGGAGGCCCTGCCT	1980
Qy	1981		CACGTGGGTCCAGCTCCCGCTCCTGTGTAGCCCATGGGGCTGCCGGCTGGCCGCCAGT	2040
Db	1981		CACGTGGGTCCAGCTCCCGCTCCTGTGTAGCCCATGGGGCTGCCGGCTGGCCGCCAGT	2040
Qy	2041		TTCTGTTGCTGCCAAAGTAAATGTGGCTCTCTGTGCCACCTGTGCTGTGAGGTGGCTA	2100
Db	2041		TTCTGTTGCTGCCAAAGTAAATGTGGCTCTCTGTGCCACCTGTGCTGTGAGGTGGCTA	2100
Qy	2101		GCTGCACAGCTGGGGCTGGGGCTCCCTCTCCTCTCTCCCACTCTTAGGGCTGCCTG	2160
Db	2101		GCTGCACAGCTGGGGCTGGGGCTCCCTCTCCTCTCTCCCACTCTTAGGGCTGCCTG	2160
Qy	2161		ACTGGAGGCCCTTCAAGGGGTTTACGTCTGGACTTATACAGGAGGCCAGAGGGCTCC	2220
Db	2161		ACTGGAGGCCCTTCAAGGGGTTTACGTCTGGACTTATACAGGAGGCCAGAGGGCTCC	2220
Qy	2221		ATGCACCTGGAAATGGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC	2280
Db	2221		ATGCACCTGGAAATGGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC	2280
Qy	2281		CTCCTAGTTGAGACACACCTAGAGAGGGTTTTTGGAGCTGTAATAAATCAGTCACCTG	2340
Db	2281		CTCCTAGTTGAGACACACCTAGAGAGGGTTTTTGGAGCTGTAATAAATCAGTCACCTG	2340
Qy	2341		GTTTTCCTCATCTTAAGCCCTTAACTGTCAGCTTCGTTTAACTAGCTCTTGATGGGAG	2400
Db	2341		GTTTTCCTCATCTTAAGCCCTTAACTGTCAGCTTCGTTTAACTAGCTCTTGATGGGAG	2400
Qy	2401		TTTCTAGGATGAACACTCTCCTCATGGATTGAAACATATGACTTATTTCTAGGGGAAGA	2460
Db	2401		TTTCTAGGATGAACACTCTCCTCATGGATTGAAACATATGACTTATTTCTAGGGGAAGA	2460
Qy	2461		GTCCTGAGGGCAACACACAGAACACAGTCCCTCAGCCACAGCACTCTCTTTTGTCT	2520











LOCUS AX429961 3410 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 100 from Patent WO0198339.  
ACCESSION AX429961  
VERSION AX429961.1 GI:21541123  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.  
TITLE Compositions and methods for the therapy and diagnosis of breast cancer  
JOURNAL Patent: WO 0198339-A 100. 27-DEC-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1..3410  
/organism="Homo sapiens"  
BASE COUNT 667 a 1014 c 945 g 783 t 1 others  
ORIGIN  
Query Match 100.0%; Score 3409.6; DB 6; Length 3410;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGAACACAGCTGCACGCGCTGGCTCCGGGTGACAGCGCGCCCTCGGCCAGGATCTGA 60  
DB 1 GGAACACAGCTGCACGCGCTGGCTCCGGGTGACAGCGCGCGCCCTCGGCCAGGATCTGA 60  
QY 61 GTGATGAGAGCTGTCCCACTGAGGTGCCCCACAGCAGAGGTGTGAGCATGGGCTGAG 120  
DB 61 GTGATGAGAGCTGTCCCACTGAGGTGCCCCACAGCAGAGGTGTGAGCATGGGCTGAG 120  
QY 121 AAGCTGACCGGACCAAGGCTGCGCAGAAATGGCGGCTGGCTGATTCCTAGGAGTT 180  
DB 121 AAGCTGACCGGACCAAGGCTGCGCAGAAATGGCGGCTGGCTGATTCCTAGGAGTT 180  
QY 181 GCGCGCAGCAAGAGGAGGCGCGAGCTTCCTGGAGCAGAGCGGAGCAAGCAGTTCTG 240  
DB 181 GCGCGCAGCAAGAGGAGGCGCGAGCTTCCTGGAGCAGAGCGGAGCAAGCAGTTCTG 240  
QY 241 GAGTGCTGAAGGCGCCCTGAGCCTTACCGCCTGGCCCACTATGTTCCAGAGGCTGTG 300  
DB 241 GAGTGCTGAAGGCGCCCTGAGCCTTACCGCCTGGCCCACTATGTTCCAGAGGCTGTG 300  
QY 301 GGTGAGCGGCTGCTGCGGACCGGAAAGCCAGCTCTTGGTGGTCAACCTGTAACCTT 360  
DB 301 GGTGAGCGGCTGCTGCGGACCGGAAAGCCAGCTCTTGGTGGTCAACCTGTAACCTT 360  
QY 361 TGGCCTGGAGGTGTGTTGGCGCAGGCAATCACCCTATGTCGCGCCTCTGCTGGAAGT 420  
DB 361 TGGCCTGGAGGTGTGTTGGCGCAGGCAATCACCCTATGTCGCGCCTCTGCTGGAAGT 420  
QY 421 GGGGGTAGAGGAGAAAGTTGATGACCAATGCTGTGGGCAATGTCAGGCTGTGGGCTGGT 480  
DB 421 GGGGGTAGAGGAGAAAGTTGATGACCAATGCTGTGGGCAATGTCAGGCTGTGGGCTGGT 480  
QY 481 CTGTGTCGCGCTCTAGGCTCAGCCAGTACCAGTGGGCTGAGCAGTATGGCGCGCGG 540  
DB 481 CTGTGTCGCGCTCTAGGCTCAGCCAGTACCAGTGGGCTGAGCAGTATGGCGCGCGG 540  
QY 541 GCCTTTCATCTGGGCACTGCTCTTGGGCATCTGCTGAGCCCTTCTCATCCCAAGGCG 600  
DB 541 GCCTTTCATCTGGGCACTGCTCTTGGGCATCTGCTGAGCCCTTCTCATCCCAAGGCG 600  
QY 601 CGGCTGGCTAGCAGGCTGCTGTGCCCCGATCCAGGCCCCCTGGAGCTGGCACTGCTCAT 660  
DB 601 CGGCTGGCTAGCAGGCTGCTGTGCCCCGATCCAGGCCCCCTGGAGCTGGCACTGCTCAT 660  
QY 661 CTTGGGCGTGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCGCT 720  
DB 661 CTTGGGCGTGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCGCT 720

QY 721 GCTCTCTGACCTCTTCGGGAGCCGACCACTGTGCCAGGCGCTACTCTGTCTATGCGCTT 780  
DB 721 GCTCTCTGACCTCTTCGGGAGCCGCGACCACTGTGCCAGGCGCTACTCTGTCTATGCGCTT 780  
QY 781 CATGATCAGTCTTGGGGCTGCTGCTGCTACTCTCTGCTGCCATTTGACTGGGACACAG 840  
DB 781 CATGATCAGTCTTGGGGCTGCTGCTGCTACTCTCTGCTGCCATTTGACTGGGACACAG 840  
QY 841 TGGCCTGGCCCCCTTACCTTGGGACCCAGGAGGAGTGCCTCTTTGGCCTGTCTACCCCTCAT 900  
DB 841 TGGCCTGGCCCCCTTACCTTGGGACCCAGGAGGAGTGCCTCTTTGGCCTGTCTACCCCTCAT 900  
QY 901 CTTCTCACCCTGCTAGCAGCACACTGCTGCTGGCTGAGGAGCAGCGCTGGGCCCCAC 960  
DB 901 CTTCTCACCCTGCTAGCAGCACACTGCTGCTGGCTGAGGAGCAGCGCTGGGCCCCAC 960  
QY 961 CGAGCCAGCAGAAGGCTGTGCGGCCCTCTCTGCTGCCGCCACTGCTGTCCATGCCGGGC 1020  
DB 961 CGAGCCAGCAGAAGGCTGTGCGGCCCTCTCTGCTGCCGCCACTGCTGTCCATGCCGGGC 1020  
QY 1021 CCGCTTGGCTTTCGGGAACCTGGGCGGCTGCTTCCCGGCTGCACAGCTGTGCTGCCG 1080  
DB 1021 CCGCTTGGCTTTCGGGAACCTGGGCGGCTGCTTCCCGGCTGCACAGCTGTGCTGCCG 1080  
QY 1081 CATGCCCGCACCTTGGCGCGCTTTCGCTGGCTGAGCTGTGCAGCTGGATGGCACATCAT 1140  
DB 1081 CATGCCCGCACCTTGGCGCGCTTTCGCTGGCTGAGCTGTGCAGCTGGATGGCACATCAT 1140  
QY 1141 GACCTTACGCTGTTTACAGGATTTGCTGGGCGAGGGGCTGTACCAGGCGCTGCCAG 1200  
DB 1141 GACCTTACGCTGTTTACAGGATTTGCTGGGCGAGGGGCTGTACCAGGCGCTGCCAG 1200  
QY 1201 AGCTGAGCGCGGCGACCGAGGCGGAGACATATGATGAAGCGCTTGGATGGCAGCCT 1260  
DB 1201 AGCTGAGCGCGGCGACCGAGGCGGAGACATATGATGAAGCGCTTGGATGGCAGCCT 1260  
QY 1261 GGGGCTGTTCTGCTGAGTGGCGCATCTCCCTGCTGCTCTCTGCTGATGAGACCGGCTGGT 1320  
DB 1261 GGGGCTGTTCTGCTGAGTGGCGCATCTCCCTGCTGCTCTCTGCTGATGAGACCGGCTGGT 1320  
QY 1321 CGAGGATTCGGCACTCGAGCAGTATATTTGGCCAGTGTGGCAGCTTTCCTGTGGCTGC 1380  
DB 1321 CGAGGATTCGGCACTCGAGCAGTATATTTGGCCAGTGTGGCAGCTTTCCTGTGGCTGC 1380  
QY 1381 CGGTGCCACATGCTGTCACAGTGTGGCGCTGGTGACAGCTTCAGCCGCGCTTCACCGG 1440  
DB 1381 CGGTGCCACATGCTGTCACAGTGTGGCGCTGGTGACAGCTTCAGCCGCGCTTCACCGG 1440  
QY 1441 GTTCACTTCTCAGCCCTGACAGATCTGCCCTACACACTGGCCCTCCCTCTACCAACCGGA 1500  
DB 1441 GTTCACTTCTCAGCCCTGACAGATCTGCCCTACACACTGGCCCTCCCTCTACCAACCGGA 1500  
QY 1501 GAAGCAGGTGTTCTGCCCCAAATACCGAGGAGACACTGGAGGTGCTAGCAGTAGGACAG 1560  
DB 1501 GAAGCAGGTGTTCTGCCCCAAATACCGAGGAGACACTGGAGGTGCTAGCAGTAGGACAG 1560  
QY 1561 CCTGATGACAGCTTCTGCGCAGGCGCTTAAGCCTTGGAGCTTCCCTTCCCTTAATGACAGCT 1620  
DB 1561 CCTGATGACAGCTTCTGCGCAGGCGCTTAAGCCTTGGAGCTTCCCTTCCCTTAATGACAGCT 1620  
QY 1621 GGGTGTGGAGGCGAGTGGCGCTGCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCTGTG 1680  
DB 1621 GGGTGTGGAGGCGAGTGGCGCTGCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCTGTG 1680  
QY 1681 TGATGTCTCCGTACGTGTGGTGGGTGAGCCCAACCGAGGCGAGGTTGCTCCGCGCGG 1740  
DB 1681 TGATGTCTCCGTACGTGTGGTGGGTGAGCCCAACCGAGGCGAGGTTGCTCCGCGCGG 1740  
QY 1741 GGGATCTGCTGACCTGCGCATCTCTGGATAGTGCCTTCTGCTGTGCCAGGTGGCCCC 1800  
DB 1741 GGGATCTGCTGACCTGCGCATCTCTGGATAGTGCCTTCTGCTGTGCCAGGTGGCCCC 1800





EVGVEEKMTWLVGIPVGLVCPVLLGSASHWRGRYRRRPFIWALSIGLILSLFL  
 IPRAGLALGLCPDRPLRLEALLILVGLLDPCGVCFPLEALLSLDRPDRHCRVA  
 YSVAFMISLGGCGLYLLPAIDWDTLSALAPVLTQEECLGLLFLILTCVAATLLVA  
 EEAALPTPEPAGLSAPSLSPHCCPCRARLAFRLNIGALLPRHOLCZEMPTLRRLFV  
 AELCSWMAIMFTFLYIDFVEGLYQVPRAPGTEARRHYDEGRMSUGLFLQCAL  
 SLVFLWMDLRVQRFTRAVYLASVAAPFAAGATCLSHSVAVYTAASAAATGFTFSAL  
 QLLPYTLASLYHREKQVFLRYRGTDTGASSEDLSMTSLFLPKPAPFPNGHVGAG  
 SGLLPPPALGASACDVSVRVVGEPTEARVPGRGICLDLIDLALDSAFLLSQVAPSL  
 FMGSTVQLSQSVTAVMSAAGLGLVAVIFATVQVFDKSLAKYSA"  
 BASE COUNT 667 a 1014 c 945 g 783 t 1 others  
 ORIGIN

Query Match 100.0%; Score 3409.6; DB 9; Length 3410;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGGAACCAAGCTGACGCGCTGGCTCCGGGTGACAGCCGCGCGCTCGGCCAGGATCGA	60
DB	1	GGGAACCAAGCTGACGCGCTGGCTCCGGGTGACAGCCGCGCGCTCGGCCAGGATCGA	60
QY	61	GTGATGAGACGTGTCCCACTGAGGTGCCCCACAGCAGCGTGTGAGCATGGCTGAG	120
DB	61	GTGATGAGACGTGTCCCACTGAGGTGCCCCACAGCAGCGTGTGAGCATGGCTGAG	120
QY	121	AAGCTGGACCGGACCAAGGGCTGGCAGAAATGGCGCCTGGCTGATTCCTAGGCAGTT	180
DB	121	AAGCTGGACCGGACCAAGGGCTGGCAGAAATGGCGCCTGGCTGATTCCTAGGCAGTT	180
QY	181	GGCGGCACAGGAGGAGCGCGCAGCTTCTGAGCAGAGCCGAGAGCAAGCAAGTCTG	240
DB	181	GGCGGCACAGGAGGAGCGCGCAGCTTCTGAGCAGAGCCGAGAGCAAGCAAGTCTG	240
QY	241	GAGTGCCTGAACGGCCCTGAGCCCTACCCGCTGGCCCACTATGTGCCAGAGGCTGTG	300
DB	241	GAGTGCCTGAACGGCCCTGAGCCCTACCCGCTGGCCCACTATGTGCCAGAGGCTGTG	300
QY	301	GGTGAAGCGCTGTGCGGCACCGGAAAGCCAGCTTGTGTGTCACACCTGTAAACCTT	360
DB	301	GGTGAAGCGCTGTGCGGCACCGGAAAGCCAGCTTGTGTGTCACACCTGTAAACCTT	360
QY	361	TGGCGTGAAGGTGTTGGCGGACGAGCATCACCTATGTGCCGCTCTGTGCTGGAGT	420
DB	361	TGGCGTGAAGGTGTTGGCGGACGAGCATCACCTATGTGCCGCTCTGTGCTGGAGT	420
QY	421	GGGGGTAGAGGAGAGTTATGACCATGTGTGGCATTTGGTCCAGTGTGGCCCTGGT	480
DB	421	GGGGGTAGAGGAGAGTTATGACCATGTGTGGCATTTGGTCCAGTGTGGCCCTGGT	480
QY	481	CTGTGTCGCTCTAGGCTCAGCCAGTGACCATGTGGCGTGGAGCGCTATGGCCGCCGCG	540
DB	481	CTGTGTCGCTCTAGGCTCAGCCAGTGACCATGTGGCGTGGAGCGCTATGGCCGCCGCG	540
QY	541	GGCCCTTATCTGGGCACGTCTTGGGCATCTGTGAGCCTCTTCTCATCCCAAGGGC	600
DB	541	GGCCCTTATCTGGGCACGTCTTGGGCATCTGTGAGCCTCTTCTCATCCCAAGGGC	600
QY	601	CGGTGGCTAGCAGGCTGTGTGCCCGGATCCAGGCCCTGGAGCTGGCACTGCTCAT	660
DB	601	CGGTGGCTAGCAGGCTGTGTGCCCGGATCCAGGCCCTGGAGCTGGCACTGCTCAT	660
QY	661	CCTGGCGTGGGCTGTGGACTTCTGTGGCAGGTGTGCTTCACTGGAGGCGCT	720
DB	661	CCTGGCGTGGGCTGTGGACTTCTGTGGCAGGTGTGCTTCACTGGAGGCGCT	720
QY	721	GCTCTCTACCTCTTCGGGACCGGACCACTGTGCCAGGCTTCTGTCTATGCTT	780
DB	721	GCTCTCTACCTCTTCGGGACCGGACCACTGTGCCAGGCTTCTGTCTATGCTT	780
QY	781	CATGATCAGTCTTGGGGCTGCTGGCTACCTCTGCTGCCATTTGACTGGGACACAG	840
DB	781	CATGATCAGTCTTGGGGCTGCTGGCTACCTCTGCTGCCATTTGACTGGGACACAG	840
QY	841	TGCCCTGGCCCCCTACCTGGGACCCAGGAGGAGTGCTCTTTGGCGCTGCTACCCCTCAT	900

DB	841	TGCCCTGGCCCCCTACCTGGGACCCAGGAGGAGTGCTCTTTGGCGCTGCTACCCCTCAT	900
QY	901	CTTCTCCTCACCCTAGCAGCCACACTGTCTGGTGTAGAGGAGCAGCGTGGGCCCCAC	960
DB	901	CTTCTCCTCACCCTAGCAGCCACACTGTCTGGTGTAGAGGAGCAGCGTGGGCCCCAC	960
QY	961	CGAGCCAGCAGAGGCGTGTGGCCCCCTCTCTTGTGCGCCCACTGCTGTCCATGCGCGGC	1020
DB	961	CGAGCCAGCAGAGGCGTGTGGCCCCCTCTCTTGTGCGCCCACTGCTGTCCATGCGCGGC	1020
QY	1021	CCGCTTGGCTTTCGGAAACCTGGCGCGCTTCCCGCGCTGCACAGCTGTGCTGCCG	1080
DB	1021	CCGCTTGGCTTTCGGAAACCTGGCGCGCTTCCCGCGCTGCACAGCTGTGCTGCCG	1080
QY	1081	CATGCCCCCGCACCTGTGCGCGCTTCTGCTGGTGTAGCTGTGAGTGTGATGCACTCAT	1140
DB	1081	CATGCCCCCGCACCTGTGCGCGCTTCTGCTGGTGTAGCTGTGAGTGTGATGCACTCAT	1140
QY	1141	GACCTTCACGCTGTTTTACACGGATTTCTGTTGGCGAGGGCTGTACAGGGCTGTGCCAG	1200
DB	1141	GACCTTCACGCTGTTTTACACGGATTTCTGTTGGCGAGGGCTGTACAGGGCTGTGCCAG	1200
QY	1201	AGCTGAGCCCGGACCGGAGGAGCAGACTATGATGAAGCGCTTCGGATGGGAGCCT	1260
DB	1201	AGCTGAGCCCGGACCGGAGGAGCAGACTATGATGAAGCGCTTCGGATGGGAGCCT	1260
QY	1261	GGGGCTGTTCTGAGTGTGCGCCATCTCCCTGCTTCTCTCTGCTGTGATGAGCGGTGT	1320
DB	1261	GGGGCTGTTCTGAGTGTGCGCCATCTCCCTGCTTCTCTCTGCTGTGATGAGCGGTGT	1320
QY	1321	GCAGCGATTCGGCAGCTCGAGCAGCTATTTGGCCAGTGTGGCAGCTTCCCTGTGCTGC	1380
DB	1321	GCAGCGATTCGGCAGCTCGAGCAGCTATTTGGCCAGTGTGGCAGCTTCCCTGTGCTGC	1380
QY	1381	CGGTGCCACATGCTGTCCACAGTGTGGCGCTGTGACAGCTTCAGCGCGCCTCACCGG	1440
DB	1381	CGGTGCCACATGCTGTCCACAGTGTGGCGCTGTGACAGCTTCAGCGCGCCTCACCGG	1440
QY	1441	GTTCACCTTCTCAGCCCTGAGATCTGCGCTTACACACTGGCCCTTCCCTTACACCGGGA	1500
DB	1441	GTTCACCTTCTCAGCCCTGAGATCTGCGCTTACACACTGGCCCTTCCCTTACACCGGGA	1500
QY	1501	GAAGCAGGTCTTCTGCCAAATACCGAGGGGACACTGGAGGTGTAGCAGTGGAGCAG	1560
DB	1501	GAAGCAGGTCTTCTGCCAAATACCGAGGGGACACTGGAGGTGTAGCAGTGGAGCAG	1560
QY	1561	CTGATGACACAGCTTCTGCCAGGCCCTAAGCCTGGAGCTTCCCTTAAATGACACCT	1620
DB	1561	CTGATGACACAGCTTCTGCCAGGCCCTAAGCCTGGAGCTTCCCTTAAATGACACCT	1620
QY	1621	GGGTGTGGAGGAGTGGCTGTCCCACTTCCACCGCGCTGTGGGGCGCTTCCCTG	1680
DB	1621	GGGTGTGGAGGAGTGGCTGTCCCACTTCCACCGCGCTGTGGGGCGCTTCCCTG	1680
QY	1681	TGATGTCTCCCTACGTGTGTGTGGTGTAGCCCAACCGAGGGCAGGTGTTCGGGGCGG	1740
DB	1681	TGATGTCTCCCTACGTGTGTGTGGTGTAGCCCAACCGAGGGCAGGTGTTCGGGGCGG	1740
QY	1741	GGGATCTGTGCTGGACCTGCCATCTCTGGATAGTGTCTTCTGCTGTCCAGGTGGCCCC	1800
DB	1741	GGGATCTGTGCTGGACCTGCCATCTCTGGATAGTGTCTTCTGCTGTCCAGGTGGCCCC	1800
QY	1801	ATCCCTGTTTATGGGCTTCCATTTGTCAGCTCAGCAGTGTGTCTGTCTATATGTGTCT	1860
DB	1801	ATCCCTGTTTATGGGCTTCCATTTGTCAGCTCAGCAGTGTGTCTGTCTATATGTGTCT	1860
QY	1861	TGCGCAGGCGCTGGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1920
DB	1861	TGCGCAGGCGCTGGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1920
QY	1921	CGACTTGGCCAAATACTCAGCGTAGAAAACCTTCCAGCACATTTGGGTGGAGGCGCTGCCT	1980
DB	1921	CGACTTGGCCAAATACTCAGCGTAGAAAACCTTCCAGCACATTTGGGTGGAGGCGCTGCCT	1980



QY 63 GATGAGAGTGTCCCACTGAGGTGCCCCACAGCAGAGGTTGAGCATGGGCTGAGAA 122  
DB 61 GATGAGAGTGTCCCACTGAGGTGCCCCACAGCAGAGGTTGAGCATGGGCTGAGAA 120  
QY 123 GCTGACCGGACCAAGAGGCTGGCAGAAATGGGCGCTGCTGCTGATTCCTAGGCAAGTGG 182  
DB 121 GCTGACCGGACCAAGAGGCTGGCAGAAATGGGCGCTGCTGCTGATTCCTAGGCAAGTGG 180  
QY 183 CGGAGCAGAGGAGAGGCGCAGCTTCTGGACGAGAGCGGAGACAGAGTTCGGA 242  
DB 181 CGGAGCAGAGGAGAGGCGCAGCTTCTGGACGAGAGCGGAGACAGAGTTCGGA 240  
QY 243 GTGCGCTGAACGCGCCCTGAGCCCTACCGCGCTGGCCCACTATGCTGACAGAGGCTGTGG 302  
DB 241 GTGCGCTGAACGCGCCCTGAGCCCTACCGCGCTGGCCCACTATGCTGACAGAGGCTGTGG 300  
QY 303 TGAGCGCGCTGTGGCGGACCGGAAAGCCAGCTTCTGCTGGTCAACCTGCTAACCTTTG 362  
DB 301 TGAGCGCGCTGTGGCGGACCGGAAAGCCAGCTTCTGCTGGTCAACCTGCTAACCTTTG 360  
QY 363 GCTTGAGGTGTGTTGGCGGACGATCCTATGTCGCGCTGCTGCTGGAGTGG 422  
DB 361 GCTTGAGGTGTGTTGGCGGACGATCCTATGTCGCGCTGCTGCTGGAGTGG 420  
QY 423 GGTGAGAGAGAGTTCATGACATGGTGTGGGCAATGGTCCAGTGTGGCGCTGTCT 482  
DB 421 GGTGAGAGAGAGTTCATGACATGGTGTGGGCAATGGTCCAGTGTGGCGCTGTCT 480  
QY 483 GTGTCCCGCTCTAGGCTCAGCCAGTACCTATGCGGTGGAGCTATGGCGCGCCCGG 542  
DB 481 GTGTCCCGCTCTAGGCTCAGCCAGTACCTATGCGGTGGAGCTATGGCGCGCCCGG 540  
QY 543 CTTTCACTGCGCACTGTCTTGGGCACTCTGCTGAGCTCTTCTCATCCCAAGGCGG 602  
DB 541 CTTTCACTGCGCACTGTCTTGGGCACTCTGCTGAGCTCTTCTCATCCCAAGGCGG 600  
QY 603 GCTGCTAGCAGGCTGTGTGCCCGGATGCCAGGCGCTGGAGCTGCGCACTGTCTATCC 662  
DB 601 GCTGCTAGCAGGCTGTGTGCCCGGATGCCAGGCGCTGGAGCTGCGCACTGTCTATCC 660  
QY 663 TGGGCTGGGCTGTGGACTTCTGTGGCCAGGTGTCTCACTCCACTGGAGGCGCTGC 722  
DB 661 TGGGCTGGGCTGTGGACTTCTGTGGCCAGGTGTCTCACTCCACTGGAGGCGCTGC 720  
QY 723 TCTCTGACCTCTTCCGGGACCGGACCTGTGCGAGGCTTCTGTCTATGCTTCA 782  
DB 721 TCTCTGACCTCTTCCGGGACCGGACCTGTGCGAGGCTTCTGTCTATGCTTCA 780  
QY 783 TGATAGTCTTGGGCTGCTGGGCTACCTCTGCTGCCATGACTGGGACACAGTG 842  
DB 781 TGATAGTCTTGGGCTGCTGGGCTACCTCTGCTGCCATGACTGGGACACAGTG 840  
QY 843 CCTTGGCGCCCTACCTGGGACCGGAGGAGTGTCTTGGCGCTGCTACCTCATCT 902  
DB 841 CCTTGGCGCCCTACCTGGGACCGGAGGAGTGTCTTGGCGCTGCTACCTCATCT 900  
QY 903 TCTTCACTGCTGAGCAGGACACACTGTGTGGTGGCTGAGGAGGAGGCTGGGCGCCACCG 962  
DB 901 TCTTCACTGCTGAGCAGGACACACTGTGTGGTGGCTGAGGAGGAGGCTGGGCGCCACCG 960  
QY 963 AGCAGCAGAGGCTGTGGGCGCCCTCTCTGTGGCGCCACCTGCTGCTGATGGGCGG 1022  
DB 961 AGCAGCAGAGGCTGTGGGCGCCCTCTCTGTGGCGCCACCTGCTGCTGATGGGCGG 1020  
QY 1023 GCTTGGCTTTCGGAACTGGGCGCCCTGCTTCCCGGCTGCACAGCTGTGTCGCGCA 1082  
DB 1021 GCTTGGCTTTCGGAACTGGGCGCCCTGCTTCCCGGCTGCACAGCTGTGTCGCGCA 1080  
QY 1083 TGCGCCGACCTCTGCGCGGCTCTTGTGGCTGAGTGTGAGCTGGATGGCACTCATGA 1142  
DB 1081 TGCGCCGACCTCTGCGCGGCTCTTGTGGCTGAGTGTGAGCTGGATGGCACTCATGA 1140  
QY 1143 CCTTCACTGCTGTTTACACGAGTTCGTGGGCGAGGCTGTACACAGGCGGCTGCCAGAG 1202

1141 CCTTCACTGCTGTTTACACGAGTTCGTGGGCGAGGCTGTACACAGGCGCTGCCAGAG 1200  
1203 CTGAGCGGCGACCGAGGCGCGGAGACACTATGATGAAGCGCTTCGATGGGAGGCTGG 1262  
1201 CTGAGCGGCGACCGAGGCGCGGAGACACTATGATGAAGCGCTTCGATGGGAGGCTGG 1260  
1263 GGTGTTCCTGTCAGTGGCGCATCTCCCTGCTTCTCTCTGTCATGAGGCGGCTGGTGC 1322  
1261 GGTGTTCCTGTCAGTGGCGCATCTCCCTGCTTCTCTCTGTCATGAGGCGGCTGGTGC 1320  
1323 AGCGATTCGCACTTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCTGTGGCTGGCG 1382  
1321 AGCGATTCGCACTTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCTGTGGCTGGCG 1380  
1383 GTGCCACATGCTTCCACAGTGTGGCGCTGTGACAGCTTCAGCGCGCTCACCGGT 1442  
1381 GTGCCACATGCTTCCACAGTGTGGCGCTGTGACAGCTTCAGCGCGCTCACCGGT 1440  
1443 TCACCTTCTCAGCGCTGCAGATCTGCTTACACACTGGCGCTCCCTTACACCGGAGAG 1502  
1441 TCACCTTCTCAGCGCTGCAGATCTGCTTACACACTGGCGCTCCCTTACACCGGAGAG 1500  
1503 AGCAGGTGTTCCTGCCCAATACCGAGGCGACACTGGAGTGTCTAGGCTGAGGACAGCC 1562  
1501 AGCAGGTGTTCCTGCCCAATACCGAGGCGACACTGGAGTGTCTAGGCTGAGGACAGCC 1560  
1563 TGATGACAGCTTCTGCCAGGCGCTTAAGCTGTGAGCTTCCCTTCCCTTAATGGACAGCTGG 1622  
1561 TGATGACAGCTTCTGCCAGGCGCTTAAGCTGTGAGCTTCCCTTCCCTTAATGGACAGCTGG 1620  
1623 GTGTGAGGCGAGTGGCTGTCTCCACCTCCACCGCGCTCTGCGGGCGCTCTGCTGTG 1682  
1621 GTGTGAGGCGAGTGGCTGTCTCCACCTCCACCGCGCTCTGCGGGCGCTCTGCTGTG 1680  
1683 ATGTCTCGTACGTGTGTGGTGTGAGCGGCGGAGGCGGAGGCTGGTTCGCGGCGG 1742  
1681 ATGTCTCGTACGTGTGTGGTGTGAGCGGCGGAGGCGGAGGCTGGTTCGCGGCGG 1740  
1743 GCATCTGCTGGACCTCGCCATCTGATAGTGTCTTCTGCTCTCCAGGCTGGCGCCAT 1802  
1741 GCATCTGCTGGACCTCGCCATCTGATAGTGTCTTCTGCTCTCCAGGCTGGCGCCAT 1800  
1803 CCTGTTCATGGGCTTCATTCAGCTCAGCAGTGTGCTGCTGCTGCTATATGCTGTG 1862  
1801 CCTGTTCATGGGCTTCATTCAGCTCAGCAGTGTGCTGCTGCTGCTATATGCTGTG 1860  
1863 CCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1922  
1861 CCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
1923 ACTTGGCCCAATACTCAGCTAGAGAACTTCCAGCAGTATGGGCTGGAGGCGCTGCTCA 1982  
1921 ACTTGGCCCAATACTCAGCTAGAGAACTTCCAGCAGTATGGGCTGGAGGCGCTGCTCA 1980  
1983 CTGGTTCACGCTCCCGCTCTGTTAGCCCGCTGCGGCGCTGCGGCGCTGCGGCGCTT 2042  
1981 CTGGTTCACGCTCCCGCTCTGTTAGCCCGCTGCGGCGCTGCGGCGCTGCGGCGCTT 2040  
2043 CTGTTGCTGCGCAAGTAAATGTTGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2102  
2041 CTGTTGCTGCGCAAGTAAATGTTGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100  
2103 TGACAGCTGGGCGCTGGGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2162  
2101 TGACAGCTGGGCGCTGGGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160  
2163 TGGAGGCTTCCAGGCGGCTTTCAGTCTGGACTTATACAGGAGGCGGAGGCTTCCAT 2222  
2161 TGGAGGCTTCCAGGCGGCTTTCAGTCTGGACTTATACAGGAGGCGGAGGCTTCCAT 2220  
2223 GCATCTGAGTTCGCGGAGCTCTGCTGAGTGGATTTACCGAGGCTCAGGCTTAACAGCTAGCT 2282

Db	2221	GC	ACTGGAA	TGGGGG	ACTG	CAGGTGG	ATTAC	CCAGGCT	CAGGGT	TAA	CAGCTAG	CCCT	228
Qy	2283	CCT	AGTTGAG	ACACAC	CACTAG	AAGAGG	TTTTGG	GAGTGA	TAA	AACTCAG	TACCTGGT	2342	
Db	2281	CCT	AGTTGAG	ACACAC	CACTAG	AAGAGG	TTTTGG	GAGCTGA	NTAA	AACTCAG	TACCTGGT	2340	
Qy	2343	TTCC	CATCTTA	AGCCCTT	AACTG	CAGCTTC	GTTTAA	TGTAG	CTT	GCATGGG	AGTT	2402	
Db	2341	TTCC	CATCTTA	AGCCCTT	AACTG	CAGCTTC	GTTTAA	TGTAG	CTT	GCATGGG	AGTT	2400	
Qy	2403	TCT	AGGATGA	AACTCT	CTCAT	GGGATTT	GAA	CATATG	-	ACT	TATTTCT	AGGGA	2460
Db	2401	TCT	AGGATGA	AACTCT	CTCAT	GGGATTT	GAA	CATATG	AA	TTTCT	AGGGA	2460	
Qy	2461	GTCT	GAGGG	CAACAC	ACAG	AGTCC	CCCTC	AGCC	CAC	AGCACT	CTCTTTT	GC	2520
Db	2461	GTCT	GAGGG	CAACAC	ACAG	AGTCC	CCCTC	AGCC	CAC	AGCACT	CTCTTTT	GC	2520
Qy	2521	GAT	CA	CCCCCT	CTTACC	TTTTAT	CAGGAT	TGG	CCCTG	TCTG	TTCCTG	CCATCA	2580
Db	2521	GAT	CA	CCCCCT	CTTACC	TTTTAT	CAGGAT	TGG	CCCTG	TCTG	TTCCTG	CCATCA	2580
Qy	2581	CAG	ACAC	AGGCA	TTTAA	TATTTAA	CTTAT	TATTTAA	CA	AA	AGTAG	AGGAATCC	2640
Db	2581	CAG	ACAC	AGGCA	TTTAA	TATTTAA	CTTAT	TATTTAA	CA	AA	AGTAG	AGGAATCC	2640
Qy	2641	TGCT	AGCTTT	TCTGT	TGGT	GTCTA	TATTTGG	TAG	GGTGGG	GAT	CCCCAAC	ATCA	2700
Db	2641	TGCT	AGCTTT	TCTGT	TGGT	GTCTA	TATTTGG	TAG	GGTGGG	GAT	CCCCAAC	ATCA	2700
Qy	2701	GGT	CCCC	TGAG	ATGCT	TGGT	CATTTGG	CTCAT	TTC	CCAGAA	TCTTCT	CCTCGGGT	2760
Db	2701	GGT	CCCC	TGAG	ATGCT	TGGT	CATTTGG	CTCAT	TTC	CCAGAA	TCTTCT	CCTCGGGT	2760
Qy	2761	CTG	CCCC	CCCA	AAATG	CGCTAAC	CCAG	CACTTGG	AAAT	TCTACT	CA	TCCAAATGATA	2820
Db	2761	CTG	CCCC	CCCA	AAATG	CGCTAAC	CCAG	CACTTGG	AAAT	TCTACT	CA	TCCAAATGATA	2820
Qy	2821	TCC	AAATGCT	TGTAC	CCCA	AGGTTAG	GTGTTGA	AGGAAG	TAG	AGG	TGGG	GCTCAGCT	2880
Db	2821	TCC	AAATGCT	TGTAC	CCCA	AGGTTAG	GTGTTGA	AGGAAG	TAG	AGG	TGGG	GCTCAGCT	2880
Qy	2881	CTC	ACGGCT	TCCCTAAC	CA	CCCTCTCT	TCTTGG	CCCA	GGCTTGG	TCC	CCCCAC	TTCCA	2940
Db	2881	CTC	ACGGCT	TCCCTAAC	CA	CCCTCTCT	TCTTGG	CCCA	GGCTTGG	TCC	CCCCAC	TTCCA	2940
Qy	2941	CTC	CCCTCT	ACTCTCT	CTAG	ACTGG	CTGATG	AA	GGCACT	GCC	CAAAATTTCC	CTACC	3000
Db	2941	CTC	CCCTCT	ACTCTCT	CTAG	ACTGG	CTGATG	AA	GGCACT	GCC	CAAAATTTCC	CTACC	3000
Qy	3001	CCC	AACTTTCC	CTAC	CCCCAA	CTTTCC	CCCA	CA	GGCTCC	CA	ACCCCTGTTTGG	AGCTACT	3060
Db	3001	CCC	AACTTTCC	CTAC	CCCCAA	CTTTCC	CCCA	CA	GGCTCC	CA	ACCCCTGTTTGG	AGCTACT	3060
Qy	3061	GC	AGGACCA	GACCA	AAAGTGG	GTTC	CCCA	AG	CCCTTGT	CCAT	CTCAG	CCCCAGAGT	3120
Db	3061	GC	AGGACCA	GACCA	AAAGTGG	GTTC	CCCA	AG	CCCTTGT	CCAT	CTCAG	CCCCAGAGT	3120
Qy	3121	AT	ATCTGT	CTGGG	AAATCTC	AC	ACAGAA	AACTC	AGG	AGC	CCCCCTG	CTCAGCTA	3180
Db	3121	AT	ATCTGT	CTGGG	AAATCTC	AC	ACAGAA	AACTC	AGG	AGC	CCCCCTG	CTCAGCTA	3180
Qy	3181	G	AGTCTAT	CTCTCA	-	GGG	GGG	GT	TAG	TGC	GT	TGCAATA	3239
Db	3181	G	AGTCTAT	CTCTCA	-	GGG	GGG	GT	TAG	TGC	GT	TGCAATA	3239
Qy	3240	TT	AG	GGGT	TGA	TATTTT	TAT	ACTG	T	AG	CAAT	CAG	3299
Db	3241	TT	AG	GGGT	TGA	TATTTT	TAT	ACTG	T	AG	CAAT	CAG	3299
Qy	3300	AAA	TTAA	AGGCT	TTTCT	TAT	3319						
Db	3301	AAA	TTAA	AGGCT	TTTCT	TAT	3320						

Query Match 85.0%; Score 2900.2; DB 9; Length 3514;  
Best Local Similarity 94.8%; Pred. No. 0;  
Matches 3155; Conservative 0; Mismatches 118; Indels 54; Gaps 13;

QY 51 CAGGATCTGAGTATGAGACGTGTCCCACTGAGGTGCCCCACAGCAGCAGGTGTGAGC 110  
DB 221 CAGGATCTGAGTATGAGATGTCTCCCACTGAGGTGCCCCACAGCAGCAGGTGTGAGC 280

QY 111 AFGGCTGAGAACTGAGACGACCCAAAGGCTGGCAGAAATGGGCGCTTGGCTGATTC 170  
DB 281 AFGGCTGAGAACTGAGACGACCCAAAGGCTGGCAGAAATGGGCGCTTGGCTGATTC 340

QY 171 CTAGGCACTTGGCGCAGCAAGGAGGAGGAGCGAGCTTCTGGAGCAGAGCCGAGAGCA 230  
DB 341 CTAGGCACTTGGCGCAGCAAGGAGGAGGAGCGAGCTTCTGGAGCAGAGCCGAGAGCA 400

QY 231 AGCAGTCTTGGAGTGCCTGAAGCGCCCTGAGCCCTACCGCCCTGGCCCACTATGTTC 290  
DB 401 AGCAGTCTTGGAGTGCCTGAAGCGCCCTGAGCCCTACCGCCCTGGCCCACTATGTTC 460

QY 291 AGAGGCTGTGGTGTAGCGCCCTGTGGGCGACCGGAAAGCCAGCTTGTGCTGTCAACC 350  
DB 461 AGAGGCTGTGGTGTAGCGCCCTGTGGGCGACCGGAAAGCCAGCTTGTGCTGTCAACC 520

QY 351 TGCTAACCTTTGGCCTGGAGGTGTCTTGGCGCAGGACATCACCTATGTGCGCCTGTGC 410  
DB 521 TGCTAACCTTTGGCCTGGAGGTGTCTTGGCGCAGGACATCACCTATGTGCGCCTGTGC 580

QY 411 TGCTGAAGTGGGGTAGAGGAGAAATTCATGACCATGTTGGTGGGCAATGGTCCAGTGC 470  
DB 581 TGCTGAAGTGGGGTAGAGGAGAAATTCATGACCATGTTGGTGGGCAATGGTCCAGTGC 640

QY 471 TGGGCTGTGTGTCTGCTCCGCTTCTAGGCTAGCCAGTACCACTGGGCTGGAGCTATG 530  
DB 641 TGGGCTGTGTGTCTGCTCCGCTTCTAGGCTAGCCAGTACCACTGGGCTGGAGCTATG 700

QY 531 GCGCGCGCGCGCCCTTCACTTGGGCACTGTCTTGGGCACTGCTGTGAGCCCTTTTCTCA 590  
DB 701 GCGCGCGCGCGCCCTTCACTTGGGCACTGTCTTGGGCACTGCTGTGAGCCCTTTTCTCA 760

QY 591 TCCCAAGGCGCGCTGTGCTAGCAGGCTGTGTGCGCGGATCCAGGCGCTTGGAGCTGG 650  
DB 761 TCCCAAGGCGCGCTGTGCTAGCAGGCTGTGTGCGCGGATCCAGGCGCTTGGAGCTGG 820

QY 651 CACTGCTCATCTGGGCTGGGCTGTGCTGAGCTTCTGTGGCAGGTGTCTCACTCCAC 710  
DB 821 CACTGCTCATCTGGGCTGGGCTGTGCTGAGCTTCTGTGGCAGGTGTCTCACTCCAC 880

QY 711 TGGAGGCGCTGTCTGTGACCTTTCCGGGACCCGACACATGTGCGCCAGGCGCTACTCTG 770  
DB 881 TGGAGGCGCTGTCTGTGACCTTTCCGGGACCCGACACATGTGCGCCAGGCGCTACTCTG 940

QY 771 TCTATGCTCATGATFAGTCTTGGGCGCTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCT 830  
DB 941 TCTATGCTCATGATFAGTCTTGGGCGCTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCT 1000

QY 831 GGGACACAGTGCCTTGGCGCCCTACTTGGGACCCAGGAGGTGCTTGTGGCTGTC 890  
DB 1001 GGGACACAGTGCCTTGGCGCCCTACTTGGGACCCAGGAGGTGCTTGTGGCTGTC 1060

QY 891 TCACCTCATCTTCTTCTCACTGTGAGCAGCCACATGTGCTGCTGAGGAGGAGGCGC 950  
DB 1061 TCACCTCATCTTCTTCTCACTGTGAGCAGCCACATGTGCTGCTGAGGAGGAGGCGC 1120

QY 951 TGGGCGCCCGAGCAGCAGAGGCTGTGGGCGCTTGTGGCGCCCTTCTTGTGGCGCTGTC 1010  
DB 1121 TGGGCGCCCGAGCAGCAGAGGCTGTGGGCGCTTGTGGCGCCCTTCTTGTGGCGCTGTC 1180

QY 1011 CATGCGCGCGCGCTTGGCTTTCGGAACTTGGGCGCTGCTTCCCGGCTGCAACGAGC 1070  
DB 1181 CGTGTGCGCGCGCGCTTGGCTTTCGGAACTTGGGCGCTGCTTCCCGGCTGCAACGAGC 1240

QY 1071 TGTGCTGCGCATGCCCCGACCCCTGCGCGGCTCTTGTGCTGAGCTGTGACGCTGA 1130  
DB 1241 TGTGCTGCGCATGCCCCGACCCCTGCGCGGCTCTTGTGCTGAGCTGTGACGCTGA 1300

QY 1131 TGGCACTCATGACCTTACGCTGTTTACACGATTTCGTGGCGAGGCGCTGTACAGG 1190  
DB 1301 TGGCACTCATGACCTTACGCTGTTTACACGATTTCGTGGCGAGGCGCTATACAGG 1360

QY 1191 GCGTGGCCAGAGCTGAGCCGCGGACCGAGGCGCGGAGACACTATGATGAGCGCTTCGGA 1250  
DB 1361 GCGTGGCCAGAGCTGAGCTGCGGACCGAGGCGCGGAGACACTATGATGAGCGCTTCGGA 1420

QY 1251 TGGCACTCTGGGCTGTTCTTCTGAGTGCCTCTCTCCCTGTGCTTCTCTCTGCTCATGG 1310  
DB 1421 TGGCACTCTGGGCTGTTCTTCTGAGTGCCTCTCTCCCTGTGCTTCTCTCTGCTCATGG 1480

QY 1311 ACCGGCTGTGAGCAGTTCGGCACTCGAGCAGTCTATTTTGGCGAGTGTGGCAGCTTTTCC 1370  
DB 1481 ACCGGCTGTGAGCAGTTCGGCACTCGAGCAGTCTATCTTGGCAGTGTGGCAGCTTTCC 1540

QY 1371 CTGTGGCTGCGGCTGCCACATGCTGTCCACAGTGTGGCGCTGTGACAGCTTCAGCCG 1430  
DB 1541 CTGTGGCTGCGGCTGCCACGCTGTCCACAGTGTGGCTGTGAGCGCTTCAGCCG 1600

QY 1431 CCTCACCGGCTTCACTTCTTCTGAGCCTGTGAGTCTTCTGCTTACACACTGCGCTTCTCT 1490  
DB 1601 CCTCACCTGCTTCACTTCTTCTGAGCCTGTGAGTCTTCTGCTTACACACTGCGCTTCTCT 1660

QY 1491 ACCACCGGAGAGCAGCTGTTCCTGCCAAATACCGAGGCGACACTGAGAGTGTCTAGCA 1550  
DB 1661 ACCACCGGAGAGCAGCTGTTCCTGCCAAATACCGAGGCGACACTGAGAGTGTCTAGCA 1720

QY 1551 GTGAGGACAGCTGATGACAGCTTCTGCGAGGCTTAAAGCTTGAAGCTTCCCTTCCCTA 1610  
DB 1721 GTGAGGACAGCTGATGACAGCTTCTGCGAGGCTTAAAGCTTGAAGCTTCCCTTCCCTA 1780

QY 1611 ATGACACGCTGGGCTGTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1670  
DB 1781 ATGACACGCTGGGCTGTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1840

QY 1671 CCTGCTGCTGTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1730  
DB 1841 CCTGCTGCTGTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1900

QY 1731 TCCGCGCGCGGCGATCTGCTGAGCCTTCCGCACTTCTGAGTGTGCTTCTGCTGCTGCTGCTG 1790  
DB 1901 TCCGCGCGCGGCGATCTGCTGAGCCTTCCGCACTTCTGAGTGTGCTTCTGCTGCTGCTGCTG 1960

QY 1791 AGGTGGCGCGGCGATCTGCTGAGCCTTCCGCACTTCTGAGTGTGCTTCTGCTGCTGCTGCTG 1850  
DB 1961 AGGTGGCGCGGCGATCTGCTGAGCCTTCCGCACTTCTGAGTGTGCTTCTGCTGCTGCTGCTG 2020

QY 1851 ATATGGTGTGCTGCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1910  
DB 2021 ATATGGTGTGCTGCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2080

QY 1911 TTGACAGAGCAGCTTGGCCAAATACTCAGCTGTGAGAACTTCCAGCAGTGTGGGCTGGA 1970  
DB 2081 TTGACAGAGCAGCTTGGCCAAATACTCAGCTGTGAGAACTTCCAGCAGTGTGGGCTGGA 2140

QY 1971 GGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2029  
DB 2141 GGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2200

QY 2030 TGGCGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2089  
DB 2201 TGGCGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2260

QY 2090 TGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2149  
DB 2261 TGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2320

QY 2150 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2209





484 TGTCCGCTCTAGGCTAGCAGTGAACACTGGCGTGGAGCTATGCGCCGCGCGCC 543  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
481 TGTCCGCTCTAGGCTAGCAGTGAACACTGGCGTGGAGCTATGCGCCGCGCGCC 540  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
544 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 603  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
541 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 600  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
604 CTGGCTAGCAGGCTGCTGTGCGCGATCCAGGCGCTGGAGCTGGAGCTGCTCATCT 663  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
601 CTGGCTAGCAGGCTGCTGTGCGCGATCCAGGCGCTGGAGCTGGAGCTGCTCATCT 660  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
664 GGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
661 GGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
724 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 783  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
721 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 780  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
784 GATCAGTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
781 GATCAGTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
844 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 903  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
841 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 900  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
904 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 963  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
901 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 960  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
964 GCGCAGAGAGGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1023  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
961 GCGCAGAGAGGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1024 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 1083  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1021 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 1080  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1084 GCGCAGAGAGGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1081 GCGCAGAGAGGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1144 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 1203  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1141 CTTTCATCTGGGCTAGTCTTGGGCTAGTCTGCTGAGCTCTTTCATCCCAAGGCGCG 1200  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1204 TGAGCGGCGACCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1239  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1201 TGAGCGGCGACCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1240 ----- 1239  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1261 AGGCTGGTGGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1240 ----- 1239  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1321 TCTCCATCTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1240 ----- 1239  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1381 CGGCTTCTGCTGGTGGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1240 ----- 1239  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1441 TAGGAGTCTGATCAGAGTGGTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1240 ----- 1239  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1501 AGTCTAGAGGAGTGGAGGATTAAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
1561 TGTCCGCTCTGCTCCGCCAACAGACTTTTCAAAATATCTACAGCGCTTCCAGCTCAGG 1620  
1240 ----- 1239

```
Db 2641 CAGCGTAGAAAATCTCCAGCACATTTGGGGTGGAGGCGCTGCCTCACTGGGTCCCAAGCTCC 2700
Qy 1998 CCCTCTCTGTAGCCCATGGGCTGCGGGCTGGCGCCAGTTCCTGTGTGCTGCCAAG 2057
Db 2701 CCGCTCTCTGTAGCCCATGGGCTGCGGGCTGGCGCCAGTTCCTGTGTGCTGCCAAG 2760
Qy 2058 TAATGTGGCTCTCTGTGTCACCCCTGTGTGCTGAGGTGGGTGAGTGCACAGCTGGGGC 2117
Db 2761 TAATGTGGCTCTCTGTGTCACCCCTGTGTGCTGAGGTGGGTGAGTGCACAGCTGGGGC 2820
Qy 2118 TGGGGCTCCCTCTCTCTCTCCCACTCTAGGGCTGCTGACGTGGAGGCTTCCAAG 2177
Db 2821 TGGGGCTCCCTCTCTCTCTCCCACTCTAGGGCTGCTGACGTGGAGGCTTCCAAG 2880
Qy 2178 GGGGTTTCAGTCTGAGCTTATACAGGAGGCGCAGAGGGCTCCAGTGCACCTGGAAATCGGG 2237
Db 2881 GGGGTTTCAGTCTGAGCTTATACAGGAGGCGCAGAGGGCTCCAGTGCACCTGGAAATCGGG 2940
Qy 2238 GACTCTGAGGTGATTACCCAGGCTCAGGGTTAAACAGCTAGGCTCTAGTTGAGACACA 2297
Db 2941 GACTCTGAGGTGATTACCCAGGCTCAGGGTTAAACAGCTAGGCTCTAGTTGAGACACA 3000
Qy 2298 CCTAGAGAGGTTTGGGAGCTGAATAAAGTCACTGCTGTTTCCCATCTCTAAGC 2357
Db 3001 CCTAGAGAGGTTTGGGAGCTGAATAAAGTCACTGCTGTTTCCCATCTCTAAGC 3060
Qy 2358 CCCTTAACCTGCAGCTTCGTTTAAATGTAGCTCTTGTGATGGAGTTCCTAGGATGAACAC 2417
Db 3061 CCCTTAACCTGCAGCTTCGTTTAAATGTAGCTCTTGTGATGGAGTTCCTAGGATGAACAC 3120
Qy 2418 TCCTCCATGGATTGAACATATG - ACTTATTTGTAGGGAAGAGTCCCTGAGGGCAAC 2475
Db 3121 TCCTCCATGGATTGAACATATGAAAGTATTTGTAGGGAAGAGTCCCTGAGGGCAAC 3180
Qy 2476 ACACAAGAACAGGTCCCTCAGCCACAGCACTGTCTTTTGTGATCCACCCCTCT 2535
Db 3181 ACACAAGAACAGGTCCCTCAGCCACAGCACTGTCTTTTGTGATCCACCCCTCT 3240
Qy 2536 TACCTTTTATCAGGATGGGCTGTGTGCTCTCTGTGGCATCAGAGACACAGGCAT 2595
Db 3241 TACCTTTTATCAGGATGT - GCCTGTGTGCTCTCTGTGGCATCAGAGACACAGGCAT 3299
Qy 2596 TTAATATTTAACTTATTTTAAAGTAGAGGAATCCATTCCTAGCTTTCTGT 2655
Db 3300 TTAATATTTAACTTATTTTAAAGTAGAGGAATCCATTCCTAGCTTTCTGT 3359
Qy 2656 GTTGTGTCTAATATTTGGGTAGGTTGGGGATCCCAACAATCAGGTCCCTGAGATAG 2715
Db 3360 GTTGTGTCTAATATTTGGGTAGGTTGGGGATCCCAACAATCAGGTCCCTGAGATAG 3419
Qy 2716 GTGGTCATGGGCTGATCATTTCCAGAAATCTTCTCTCTGGGCTGCGCCCAAAAT 2775
Db 3420 GTGGTCATGGGCTGATCATTTCCAGAAATCTTCTCTCTGGGCTGCGCCCAAAAT 3479
Qy 2776 GCTTAACCCAGGACCTGGAATTTCTACTATCCAAATGATAATCCAAATGCTGTAC 2835
Db 3480 GCTTAACCCAGGACCTGGAATTTCTACTATCCAAATGATAATCCAAATGCTGTAC 3539
Qy 2836 CCAAGGTTAGGTTGTAAGGAAGGTAGAGGTGGGCTTCAGGTCTCAACGGCTTCCT 2895
Db 3540 CCAAGGTTAGGTTGTAAGGAAGGTAGAGGTGGGCTTCAGGTCTCAACGGCTTCCT 3599
Qy 2896 AACACACCTCTTCTCTTGGCCAGCCTGTTCCCGCCACTTCCACTCCCTCTACTCTC 2955
Db 3600 AACACACCTCTTCTCTTGGCCAGCCTGTTCCCGCCACTTCCACTCCCTCTACTCTC 3659
Qy 2956 TCTAGAGCTGGCTGATGAAGCACTGCCCAAAATTTCCCTTACCCCAACTTTCCCTTA 3015
Db 3660 TCTAGAGCTGGCTGATGAAGCACTGCCCAAAATTTCCCTTACCCCAACTTTCCCTTA 3719
Qy 3016 CCCCCAACCTTTCCCAACAGCTCCCAACCCCTGTTTGGAGCTACTGACGAGCACAAGCA 3075
Db 3720 CCCCCAACCTTTCCCAACAGCTCCCAACCCCTGTTTGGAGCTACTGACGAGCACAAGCA 3779
```

```
Qy 3076 CAAAGTGGGGTTTCCCAAGCCCTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTGGG 3135
Db 3780 CAAAGTGGGGTTTCCCAAGCCCTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTGGG 3839
Qy 3136 GAATCTCACAGAAACTCAGGAGCACCCCTGCTGAGCTAAGGAGGCTTATCTCTC 3195
Db 3840 GAATCTCACAGAAACTCAGGAGCACCCCTGCTGAGCTAAGGAGGCTTATCTCTC 3899
Qy 3196 AGGGGGGTTTAAAGTCCGCTTTGCAATATGTCGCTTATTTATTTAGCGGGTGAATAT 3255
Db 3900 AGGGGGGTTTAAAGTCCGCTTTGCAATATGTCGCTTATTTATTTAGCGGGTGAATAT 3959
Qy 3256 TTTTACTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3315
Db 3960 TTTTACTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4019
Qy 3316 TTTATATGTTTAAAAA 3330
Db 4020 TTTATATGTTTAAAAA 4034

RESULT 10
AX267730
LOCUS AX267730 4034 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 704 from Patent WO0173032.
ACCESSION AX267730
VERSION AX267730.1 GI:16516402
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,
Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H.,
Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A., Hepler, W.T.
and Henderson, R.A.
TITLE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0173032-A 704 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source 1. 4034
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 721 a 1226 c 1141 g 946 t
ORIGIN
Query Match 75.8%; Score 2585.4; DB 6; Length 4034;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 3325; Conservative 0; Mismatches 1; Indels 709; Gaps 3;
Qy 4 AACACGCTGCACGCGCTCCGGTGCAGCCGCTGACAGCCCGCGCTCGGCCAGGATCTGAGTG 63
Db 1 AACACGCTGCACGCGCTCCGGTGCAGCCGCTGACAGCCCGCGCTCGGCCAGGATCTGAGTG 60
Qy 64 ATGAGAGCTGTCCTCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGCTGAGAAG 123
Db 61 ATGAGAGCTGTCCTCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGCTGAGAAG 120
Qy 124 CTGACCGCGCACCAAGGGCTGGCAGAAATGGCGGCTGCTGATTTCTTAGGCAGTTGCG 183
Db 121 CTGACCGCGCACCAAGGGCTGGCAGAAATGGCGGCTGCTGATTTCTTAGGCAGTTGCG 180
Qy 184 GGCAGCAAGGAGGAGGCGCGAGCTTCTGGAGCAGCGGAGCAGCAAGTCTTCTGGAG 243
Db 181 GGCAGCAAGGAGGAGGCGCGAGCTTCTGGAGCAGCGGAGCAGCAAGTCTTCTGGAG 240
Qy 244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCTGCGCCACTATGTCAGAGGCTGTGGGT 303
Db 241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCTGCGCCACTATGTCAGAGGCTGTGGGT 300
```

QY 304 GAGCGCCTGCTGCGGCACCGGAAAGCCAGCTCTTGTGGTCAACCTGCTAACTTTGG 363  
Db 301 GAGCGCCTGCTGCGGCACCGGAAAGCCAGCTCTTGTGGTCAACCTGCTAACTTTGG 360  
QY 364 CTTGAGAGTGTGTTTGGCGGAGGAGATCACCTATGTCGGCTGCTGCTGGAAGTGG 423  
Db 361 CTTGAGAGTGTGTTTGGCGGAGGAGATCACCTATGTCGGCTGCTGCTGGAAGTGG 420  
QY 424 GGTAGAGGAGAGTTCATGACCATGCTGGGCATTTGTCAGTGTGGCTGCTGCTG 483  
Db 421 GGTAGAGGAGAGTTCATGACCATGCTGGGCATTTGTCAGTGTGGCTGCTGCTG 480  
QY 484 TGTCCGCTCTTGTAGGCTCAGCAGTACCTGAGCTGAGCTATGTCGGCTGCTGCTG 543  
Db 481 TGTCCGCTCTTGTAGGCTCAGCAGTACCTGAGCTGAGCTATGTCGGCTGCTGCTG 540  
QY 544 CTTGATCTGGGACTGCTTGTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTG 603  
Db 541 CTTGATCTGGGACTGCTTGTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTG 600  
QY 604 CTGGCTAGCAGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTG 663  
Db 601 CTGGCTAGCAGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTG 660  
QY 664 GGGCTGGGCTGCTGAGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTG 723  
Db 661 GGGCTGGGCTGCTGAGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTG 720  
QY 724 CTCTGACCTCTCCGGGACCGGACCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTG 783  
Db 721 CTCTGACCTCTCCGGGACCGGACCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTG 780  
QY 784 GATCAGTCTTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTG 843  
Db 781 GATCAGTCTTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTG 840  
QY 844 CTTGCGCCTTACCTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCT 903  
Db 841 CTTGCGCCTTACCTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCT 900  
QY 904 CTTGACCTGCTGAGCAGCAGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCT 963  
Db 901 CTTGACCTGCTGAGCAGCAGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCT 960  
QY 964 GCGAGCAGAGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTGAG 1023  
Db 961 GCGAGCAGAGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGCTGAG 1020  
QY 1024 CTTGCTTTCGGGAACTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCT 1083  
Db 1021 CTTGCTTTCGGGAACTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCT 1080  
QY 1084 GCGCGGACCTGCGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGAG 1143  
Db 1081 GCGCGGACCTGCGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGCTGAGCTGCTGAG 1140  
QY 1144 CTTGAGCTGCTTACAGGATTTGCTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGAG 1203  
Db 1141 CTTGAGCTGCTTACAGGATTTGCTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGAG 1200  
QY 1204 TGAGCGGCGGACCGGCTGCTGCTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGAG 1239  
Db 1201 TGAGCGGCGGACCGGCTGCTGCTGGGCTGCTGCTGGGCTGCTGCTGAGCTGCTGAG 1260  
QY 1240 ----- 1239  
Db 1261 AGGTGCTGGGAGCGGCGGACAGACACTCGGGCTGCTGCTGGGCTGCTGCTGAGCTG 1320  
QY 1240 ----- 1239  
Db 1321 TCTCCATCTGGCGGCTGCTGCTGAGGAGTGGGCTGAGCGGCTGCTGCTGAGCTG 1380  
QY 1240 ----- 1239

Db 1381 CGGCTTCTCATGGGTGGGAACATCTCTGCTTGGGTTTCAGGAAGGCTCTTGGCTGCTC 1440  
QY 1240 ----- 1239  
Db 1441 TAGGAGTCTGATCAGAGTCTGTTGCCAGTTTGACAGAGAAAGGGGAGCTTATTCAA 1500  
QY 1240 ----- 1239  
Db 1501 AGTCTAGAGGAGTGGAGAGTTAAGGCTTGAATTCAGATCTGCTGTTTCCAGCGCAG 1560  
QY 1240 ----- 1239  
Db 1561 TGTGCTCTGCTCTCCCAACAGCTTTCACAAATTAATCTCACCAGCGCTTTCAGCTCAGG 1620  
QY 1240 ----- 1239  
Db 1621 CGTCTAGAAAGCTTGAAGCCTATGGCCAGCTGCTTGTGTTTCCCTCTCACCAGCT 1680  
QY 1240 ----- 1239  
Db 1681 GTCTCTACAGCTGAGACTCCAGGAAACCTTTCAGACTTACCTTCTGCTTCCAGCAAG 1740  
QY 1240 ----- 1239  
Db 1741 GCGCTTGGCCACATTTCTCTGAGGCTCAGTGAAGAAACCTTTCAGCTTCCCTTGTAGAGGTA 1800  
QY 1240 ----- 1239  
Db 1801 GAAAGGGAAGGCTGCTGGGAGCAGGCTGGTCCACAGCAGCTCTCGTGCAGCAGTAC 1860  
QY 1240 ----- 1239  
Db 1861 CTGTGTTCCGCTTCTCTATCTCCCTGAGACTGCTCCGACCTTCCCTCCCAGGCTCTGT 1920  
QY 1240 ----- 1239  
Db 1921 CTGATGGCCCTCTCCCTCTGTCAGGCTTCCGATGGGCTGCTGCTGCTGCTGCTGCT 1980  
QY 1278 GCGCATCTCTCTGCTGCT 1337  
Db 1981 GCGCATCTCTCTGCTGCT 2040  
QY 1338 GAGCAGTCTATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1397  
Db 2041 GAGCAGTCTATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100  
QY 1398 CCACAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1457  
Db 2101 CCACAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
QY 1458 TGCAGATCT 1517  
Db 2161 TGCAGATCT 2220  
QY 1518 CCAATACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1577  
Db 2221 CCAATACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
QY 1578 TGCAGGCTCTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1637  
Db 2281 TGCAGGCTCTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340  
QY 1638 GCTGCT 1697  
Db 2341 GCTGCT 2400  
QY 1698 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1757  
Db 2401 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
QY 1758 TGCATCTCTGATAGTCT 1817



were constructed by oligo-capping method  
(Sugano et al., Institute of Medical Science, University of  
Tokyo).

Custom primer used for sequencing  
( 5' end primer [CTTGTCTCTAAAGCTGCG];  
3' end primer [CGACTGCTGAGCAGACA] ).

## FEATURES

## source

1. .2917  
/organism="Macaca fascicularis"  
/db\_xref="taxon:9541"  
/clone="QmoA-10594"  
/sex="male"  
/tissue\_type="medulla oblongata"  
/clone\_lib="macaque brain cDNA library QmoA"  
/dev\_stage="adult"  
24. .1529

## CDS

/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAB60745.1"  
/db\_xref="GI:14388391"

/translation="MTWVIGIPVLVSLVLLGSNDHWRGRRPFFIWLISGLI  
LISLFLIPAGHLAGLCPDRPLELALLIGVGLDFCGQVCFTPLLEALLSDFRDP  
DUCRQAYSVTWISLGGCGYLLFAIDMTSALAPYLGTOEECLFLGLTLFLFCVA  
ATLLVAEALPAPAEGLSAPLSHCCPCWRLAFNGLALLPRLHQLCCMRPT  
LRLFLVAELCSWALMTFTLYDFVGEGLYQVPAELGTEARHDEYGVRRMSLGL  
FLQCAISLVSLVORLVQFTRAVLASVAAPFAAGATCLSHSAVAVTASALGL  
FTFSALIDLPYTLASLYHREOVFLPKYRGDAGGTSSRDSLMTSLPKPKGAPPNQ  
HYVAGSGLLPPLPACGASACDVSVRVVVEPTEARVVPGRGICLDLALDSAPLLS  
QVAPSLFMSISVQLSQSVTAYMWSAGLGLVLIYFATQVDFDKSLAKYSV"

BASE COUNT 514 a 868 c 815 g 720 t  
ORIGIN

Query Match 74.7%; Score 2545.6; DB 9; Length 2917;  
Best Local Similarity 94.6%; Pred. No. 0;  
Matches 2792; Conservative 0; Mismatches 104; Indels 54; Gaps 13;

QY 417 AAGTGGGGTAGAGAGAGTTCATGACCATGGTGGGCAATGGTCCAGTGGGCC 476  
DB 1 AAGTGGGGTAGAGAGAGTTCATGACCATGGTGGGCAATGGTCCAGTGGGCC 60  
QY 477 TGTCTGTCTCCCGTCTAGGCTCAGCCAGTGGAGTGGGCGTGGAGCGCTATGGCCGCC 536  
DB 61 TGTCTGTCTCCCGTCTAGGCTCAGCCAGTGGAGTGGGCGTGGAGCGCTATGGCCGCC 120  
QY 537 GCGGCGCTCTATCTGGGCACTGTCTTGGGCACTGTCTGAGCCCTCTTCTCATCCCAA 596  
DB 121 GCGGCGCTCTATCTGGGCACTGTCTTGGGCACTGTCTGAGCCCTCTTCTCATCCCAA 180  
QY 597 GCGGCGCTGGCTAGAGGGCTGTGTGCCCGGATCCAGGCCCTGGAGCTGGCACTGC 656  
DB 181 GCGGCTGGCTGGGCGTGTGTGCCCGGATCCAGGCCCTGGAGCTGGCACTGC 240  
QY 657 TCATCTGGGCTGGGCGTGTGTGTGGGCACTGTGTGGGCACTGTGTGGGCACTGTGTGGG 716  
DB 241 TCATCTGGGCTGGGCGTGTGTGTGGGCACTGTGTGGGCACTGTGTGGGCACTGTGTGGG 300  
QY 717 CCCTGCTCTCTGACCTCTTCCGGGACCCGAGCACTGTGTGGGCACTGTGTGGGCACTGTGTGGG 776  
DB 301 CCCTGCTCTCTGACCTCTTCCGGGACCCGAGCACTGTGTGGGCACTGTGTGGGCACTGTGTGGG 360  
QY 777 CCTTCATGATGAGTCTTGGGCGCTGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 836  
DB 361 CCTTCATGATGAGTCTTGGGCGCTGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCG 420  
QY 837 CCAGTGGCTGGGCGCTGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGG 896  
DB 421 CCAGTGGCTGGGCGCTGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGG 480  
QY 897 TCATCTCTCTACCTCTGGCTAGAGCCACATGCTGGTGGCTGAGGAGGCGAGGCTGGGCGTGGGCG 956  
DB 481 TCATCTCTCTACCTCTGGCTAGAGCCACATGCTGGTGGCTGAGGAGGCGAGGCTGGGCGTGGGCG 540  
QY 957 CCACCGAGGCGACAGAGGGCTGTGGGCGGCTGTGGGCGGCTGTGGGCGGCTGTGGGCGGCTGTGGG 1016

DB 541 CGCGCGAGCCAGCGAAGGGCTGTGTGCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 600  
QY 1017 GGGCGCGCTTGGCTTTCGGAACCTTGGCGCCCTGCTTCCCGGCTGACACAGCTGTGCT 1076  
DB 601 GGGCGCGCTTGGCTTTCGGAACCTTGGCGCCCTGCTTCCCGGCTGACACAGCTGTGCT 660  
QY 1077 GCGCATGCCCGCGCACCTTGCCTGGCTGCTTTCGCTGGCTGAGCTGTGACAGCTGGATGGCAC 1136  
DB 661 GCGCATGCCCGCGCACCTTGCCTGGCTGCTTTCGCTGGCTGAGCTGTGACAGCTGGATGGCAC 720  
QY 1137 TCATGACCTTACAGCTGTGTTTACAGGATTTGCTGGGCGAGGGCTGTGACAGGCGCTGC 1196  
DB 721 TCATGACCTTACAGCTGTGTTTACAGGATTTGCTGGGCGAGGGCTGTGACAGGCGCTGC 780  
QY 1197 CCAGAGCTGAGCGCGGCGACCGAGGCGCGGAGACACTATGATGAAGGGCTTCGGATGGCA 1256  
DB 781 CCAGAGCTGAGCTGGGCGACCGAGGCGCGGAGACACTATGATGAAGGGCTTCGGATGGCA 840  
QY 1257 GCCTGGGCTGTCTTGCAGTGGCGCATCTCCCTGGCTTCTCTGTGTGATGACACCGGC 1316  
DB 841 GTCTGGGCTGTCTTGCAGTGGCGCATCTCCCTGGCTTCTCTGTGTGATGACACCGGC 900  
QY 1317 TGGTGACGAGTTCGGCACTGACAGCTATTTGGCCAGTGTGACAGCTTTCCTCTGG 1376  
DB 901 TGGTGACGAGTTCGGCACTGACAGCTATCTTGGCCAGTGTGACAGCTTTCCTCTGG 960  
QY 1377 CTGCGGCTGCGACATGCTGTCCACAGTGTGGCGCTGGTGACAGCTTCAGCGCCCTCA 1436  
DB 961 CTGCGGCTGCGACATGCTGTCCACAGTGTGGCTGTGGTGACGCTTCAGCGCCCTCA 1020  
QY 1437 CCGGCTTACCTTCTCAGCCCTGACAGTCTGCGCCCTACACACTGGCCCTCCCTCTACACAC 1496  
DB 1021 CTGCTTACCTTCTCAGCCCTGACAGTCTGCGCCCTACACACTGGCCCTCCCTCTACACAC 1080  
QY 1497 GGGAGAGCAGTGTCTCTGCCCAAATACCGAGGGGACACTGGAGGTGTAGCAGTGGAG 1556  
DB 1081 GGGAGAGCAGTGTCTCTGCCCAAATACCGAGGGGACACTGGAGGTGTAGCAGTGGAG 1140  
QY 1557 ACAGCTGATGACAGCTTCTTGCAGGCCCTAAGCCTGGAGCTTCCCTCTAATGGAC 1616  
DB 1141 ACAGCTGATGACAGCTTCTTGCAGGCCCTAAGCCTGGAGCTTCCCTCTAATGGAC 1200  
QY 1617 ACGTGGGTGTGGAGGAGTGGCTGTCCACCTCCACCGCCCTCTGCGGGGCTCTG 1676  
DB 1201 ACGTGGGTGTGGAGGAGTGGCTGTCCCGCTCTCCACCGCCCTCTGCGGGGCTCTG 1260  
QY 1677 CCTGTGTCTCTCCGTACGTGTGGTGGGTGAGCCACCGAGGCCAGGCTGGTCCGG 1736  
DB 1261 CCTGTGTCTCTCCGTACGTGTGGTGGGTGAGCCACCGAGGCCAGGCTGGTCCGG 1320  
QY 1737 GCGGGGCACTGCTGCGACCTCCCATCTCGGATAGTGCCTTCTGCTGCCAGGTGG 1796  
DB 1321 GCGGGGCACTGCTGCGACCTCCCATCTCGGATAGTGCCTTCTGCTGCCAGGTGG 1380  
QY 1797 CCCCATCCCTGTTTATGGGTCTCATGTGTCAGCTCAGCCAGTGTGTCACGTCTATATGG 1856  
DB 1381 CCCCATCCCTGTTTATGGGTCTCATGTGTCAGCTCAGCCAGTGTGTCACGTCTATATGG 1440  
QY 1857 TGTCTGCGCAGGCTGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1916  
DB 1441 TGTCTGTCAGGCTGGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1500  
QY 1917 AGAGGAGTGGGCAAAATCTACGCTAGAAAATCTCCAGCACATTTGGGGTGGAGGCTCT 1976  
DB 1501 AGAGGAGTGGGCAAAATCTACGCTAGAAAATCTCCAGCACATTTGGGGTGGAGGCTCT 1560  
QY 1977 GCCTACCTGGGTCCAGCTCCCGCTC- CTGTTAGCCCCATGGGGTGGCGGCTGGCGG 2035  
DB 1561 GCCTACCTGGGTCCAGCTCCCGCTCTTTGTTAGCCCCATGGGGTGGCGGCTGGCGG 1620  
QY 2036 CCAAGTTCTGTGTGCCAAAGTAATGTGGCTCTCTGCTGCCACCTCTGCTGCTGCTGCTGCTGCT 2095

Db 1621 CCAAGTTCTGTGTCGCAAGTAATGTGGCTCTCTGCTGCCACCCCTGTGCTGCTGAGGT 1680  
QY 2096 GCGTAGCTGCACAGCTGGGGCTGGGGCTCCCTCTCTCTCTCCCGAGCTCTTAGGGCT 2155  
Db 1681 GCGTAGCTGCACAGCTGGGGCTGGGGCTCCCTCTCTCTCCCGAGCTCTTAGGGCT 1740  
QY 2156 GCCTGACTGGAGGCTTCCAGAGGGGTTTCACTGCTGAGCTTATACAGGGAGGCGAGAAG 2215  
Db 1741 GCCTGACTGGAGGCTTCCAGAGGGGTTTCACTGCTGAGCTTCTACAGGGAGGCTAGAAG 1800  
QY 2216 -----GCTCCATGCACCTGGAAATCGGGGACTCTGAGGTGGATTACCCA 2259  
Db 1801 CGAGGGCATTTGATTCGCTCCATGCACCTGGAAATGTGGGACTCTGAGGTGGATTACCCA 1860  
QY 2260 GCCTCAGGGTTAAGAGCTAGCTCTCTAGTTGAGACACACTAGAGAGGGTTTGGGAG 2319  
Db 1861 GCTCAGGGTTAAGAGCTAGCTCTCTAGTTGAGACACACTAGAGAGGGTTTGGGAG 1920  
QY 2320 CTGAATAACCTAGCTAGCTGGTTTCCCATCTCTAAG--CCGCTTAACTCGAGCTTCGTT 2378  
Db 1921 CTGAGTAACTAGCTAGCTGGTTTCCCATCTCTAAGCCCTTAACTCGAGCTTCGTT 1980  
QY 2379 TAATCTAGCTCTGCTGATGAGGTTTCTAGGATGAACACTCTCTCATGGGATTGAACAT 2438  
Db 1981 TAATCTAGCTCTGCTGATGAGGTTTCTAGGATGAACACTCTCTCATGGGATTGAACAT 2040  
QY 2439 ATG--ACTTATTTAGGGAAGAGTCTCTGAGGGGCAACACAGAACAGCTCCCTC 2496  
Db 2041 ATGAAAGTTATTTAGGGAAGAGTCTCTGAGGGCAACAC-----ACCAGGTCCCTC 2094  
QY 2497 AGCCACAGCACTGCTCTTTGCTGATCCACCCCTCTTACCTTTTATCAGGATGGC 2556  
Db 2095 AGCCACAGCACTGCTCTTTGCTGATCC--CCGACTCTTACCTTTTATCAGGACGTGC 2153  
QY 2557 CTGTTGGTCTCTCTGCTGATCAGAGACAGGCAATTTAAATATTAACTTATTAT 2616  
Db 2154 CTATTTGGTCTCTCTGCTGATCAGAGACAGGCAATTTAAATATTAACTTATTAT 2213  
QY 2617 TTAACAAAGTAAAGGGAATCAATGCTGATGCTTTCTGTTGTTGCTAATATTGGGT 2676  
Db 2214 TTAACAAAGTAAAGGGAATCAATGCTGATGCTTTCTGTTGTTGCTAATATTGGGT 2273  
QY 2677 AGGTGGGGATCCCAACATCAGTCCCTGATAGCTGCTGATGCTGCTGCTCATATT 2736  
Db 2274 AGGT--GGGATCCCAACATCAGTCCCTGATAGCTGCTGATGCTGCTGCTCATATT 2331  
QY 2737 GCCAGAAATCTTCTCTCTGGGCTGCTGGCCGCCCAAAATGCTTAACCCAGACCTTGGA 2796  
Db 2332 GCCAGAAATCTTCTCTCTGGGCTGCTGGCTCTCAAAATGCTTAACCCAGACCTTGGA 2391  
QY 2797 ATTCTACTATCCCAATGATAATTTCCAAATGCTGTTACCAAGGTTAGGTTGAAG 2856  
Db 2392 ATTCTACTATCCCAATGATAATTTCCAAATGCTGTTACCAAGGTTAGGTTGAAG 2451  
QY 2857 AGGTGAGGGTGGGGCTCAGTCTCAAGGCTTCCCTAACCCCTCTTCTCTTCTGGC 2916  
Db 2452 AGGTGAGGGTGGGGCTCAGTCTCAAGGCTTCCCTAACCCCTCTTCTCTTCTGGC 2511  
QY 2917 CCAGCTGCTGTTCCCGCCACTTCCACTCCCTCTACTCTCTCTAGGACTGGGCTGATGAAG 2976  
Db 2512 CCAGCTGCTGTTCCCGCCACTTCTACTCCCTCTACTGCTCTAGGACTGGGCTGATGAAG 2571  
QY 2977 GCATGCCCAAAATTTCCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCT 3036  
Db 2572 GCATGCCCTGAAATTTCCCT-----ACCCCAACTTTCCCACTGGC 2614  
QY 3037 TCCACAACTCTTTGGAGCTACTGCAGACAGAGCAGAGGCTGCTTCCCAAGCC 3096  
Db 2615 TCCACAACTCTTTGGAGCTTTGCAGACAGAGCAGAGGCTGCTTCCCAAGCC 2674  
QY 3097 TTTGTCATCTCAG--CCCCAGAGTATATCTGCTGTTGGGGAATCTACACAGAACTCA 3155  
Db 2675 TTTGTCATCTCAGCCCCCAGAGTATATCTGCTGTTGGGGAATCTACACAGAACTCA 2734

QY 3156 GGAGCACCCCTGCTGAGCTAAGGAGGCTCTTATCTCTCAGGGGGGTTTAAAGTCCGT 3215  
Db 2735 GGAGCACCCCTGCTGAGCTAAGGAGGCTCTTATCTCTCA--GGGGGGTTTAAAGTCCGT 2792  
QY 3216 TTGCAATAATGCTGCTTATTTATTTAGCGGGTGAATATTTATCTCTAAGTGAAGCA 3275  
Db 2793 TTGCAATAAT--GTCTTATTTATTTAGCGGGCAATATTTATCTCTAAGTGAAGCA 2849  
QY 3276 TCAGAGTAAATGCTTATGTCACAAAATTAAGGCTTCTTATATGTTTAAAAA 3335  
Db 2850 TC--ACTATAATGCTTATGTCATGAATTAAGGCTTCTTATATGTTTAAAAA 2907  
QY 3336 AAAAAAAA 3345  
Db 2908 AAAAAAAA 2917

RESULT 12  
AX200994  
LOCUS  
DEFINITION Sequence 624 from Patent WO0151633.  
ACCESSION AX200994  
VERSION AX200994.1 GI:15390821  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2904)  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,  
Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.  
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer  
JOURNAL Patent: WO 0151633-A 624 19-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1. 2904  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 542 a 875 c 773 g 714 t  
ORIGIN

Query Match 64.4%; Score 2196.4; DB 6; Length 2904;  
Best Local Similarity 88.2%; Pred. No. 0;  
Matches 2560; Conservative 0; Mismatches 1; Indels 343; Gaps 2;

QY 770 GTCTATGCCCTCATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGAC 829  
Db 1 GTCTATGCCCTCATGATCAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGAC 60  
QY 830 TGGGACACCAAGTCCCTGCGCCCTACCTGGGACCCAGAGAGTGCCTCTTTGGCGCTG 899  
Db 61 TGGGACACCAAGTCCCTGCGCCCTACCTGGGACCCAGAGAGTGCCTCTTTGGCGCTG 120  
QY 890 CTACCCCTCATCTCTCAGCTGCTAGCAGCCACACTGCTGCTGCTGAGGAGGAGCG 949  
Db 121 CTACCCCTCATCTCTCAGCTGCTAGCAGCCACACTGCTGCTGCTGAGGAGGAGCG 180  
QY 950 CTGGGCCCCCAGGACGAGCAGAGGCTGCTGGGCCCCCTCTTGTGCGCCCACTGCTGT 1009  
Db 181 CTGGGCCCCCAGGACGAGCAGAGGCTGCTGGGCCCCCTCTTGTGCGCCCACTGCTGT 240  
QY 1010 CCATGCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTCCCGGGCTGCACAG 1069  
Db 241 CCATGCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTCCCGGGCTGCACAG 300  
QY 1070 CTGTGCTGCGGCATGCCCCCGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1129  
Db 301 CTGTGCTGCGGCATGCCCCCGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 1130 ATGGCACTCATGACCTTTCAGCTGCTTTTACACGAGATTTCTGTTGGGCGAGGGGCTGACAG 1189



338	AGAGCTGGCGTTTCGGT---	GCCGCACAGCGCGCTCA-----	CCG	300	
716	ProProGlyAlaGlnGlyProAlaGlyProAlaGlyProGlyGlySerArgGlyAspProGlyProGlyPro			735	
299	ACAGCCTCTGGACCATAGTGGCGCCAGCGGG-----			269	
736	GlyAlaGlnGlyPro-AlaGlyProGlyGlySerArgGlyAspProGlyProGlyProGlyAl			755	
268	-----TAGGGCTCAGGGCGCGCTTCAGGCACATCCAGAACTGCTTCGTCGCGCTC			219	
755	agInGlyProAlaGlyProGlyGlySerArgGlyAspPro-----			768	
218	TCGTCACAAGCTGGCGCTCTCT-----	CCTTGCTCGCGCAACTGCCTAGGAATCA	165		
769	-GlyProProGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyPro			788	
164	GCCAGGCGCCCATTC--	TGCCAGCCCTTTGGTGCC--	CGTTCAGCTTCT--	CAGCC	114
788	oLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLys-G				808
113	CATGCTCAACACCTCTGCTGTGGGGACCTCACTAGTGGGACACGCTCTCATCTACCTCAGATC			54	
808	lyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspPro-----			GlyProp	825
53	CTGGCCGAGGCGCGGCTGTCA	CCGGAGCCAGC	19		
825	roglyAlaGlnGlyProAlaGlyProGlyGlySer				836

US-09-759-143-110 (1-3410) x US-08-931-820-4 (1-1057)

QY	3051	AAACAGGGTGTGGAGCTGTGGGAAAGTTGGGGTAGGGAAAGTTGGGGTAGGGGA	2992
Db	13	LysSerGlyValAlaValGlyGlyLeuAlaGlyTyrProGlyProAlaGlyProGlyGly	32
QY	2991	AAATTTGGGAGTGCCCTTCATCAGCCAGCTCTAGAGAGATAGAGGGG	2943
Db	33	-----ProProGlyPro-ProGlyThrSerGlyHisProGlySer	45
QY	2942	-AGTGAAGTGGGGAAACACAGCTGGCCCAAGAGAGAGGGTGTAGGGAAGCCGTT	2884
Db	45	rProGlySerProGlyTyrGlnGlyProGlyGlyGluProGlyGlyGlnAlaGly	64
QY	2883	GAGACCTGAAGCCGCCCTCTACCTTCCTTCAACACCCCTTAACCTTGGGTAAACGCA	2824
Db	64	rGlyProProGlyProGly	71
QY	2823	GGAAATTATCATTTGGGATGAGTAGAATTTCCAAGTGCTGGTTAGGCATTTTGGGGGC	2764
Db	72	-----AlaIleGlyProSerGlyProAlaGlyLysAspGlyGluSerGlyArgProGly	89
QY	2763	CAGACCCAGGGAAGAAGATTCTGCGCAATGATCAGCCCAATCAGCAGCTATCTCAGGGG	2704
Db	89	yArgProGlyGluArgGlyLeu	104
QY	2703	ACCTGATTTGGGGATCCCCACCCTACCCAAATATTAGACACCAACACAGA-	2651
Db	104	yProAla-----GlyIle-ProGlyPheProGlyMetLys--GlyHisArgGlyPheA	121
QY	2650	-----AAGCTAGCAATGGATTCCCTTCTACTTTGTTTAAATAAATAAG	2608
Db	121	spGlyArgAsnGlyGluLysGlyGluThrGlyAlaProGly	135
QY	2607	TTAAATATTAAATGCTGTGTCGTGATGCCAACAGAAAGCAACAGGCCACATCC	2548
Db	135	eulysGlyGluAsnGlyLeuProGlyGluAsnGlyAlaProGlyProMetGlyProArg	155
QY	2547	TGATAAAGGTAAAGGGGGTGGATCAGCAAAAGACAGTGTGTGGGTGAGGGGACC	2488
Db	155	lyAlaProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsn	174
QY	2487	TGGTTCTTGTTGTTCGCCCTCAGGACTCTTCCCTACAAATAAGTCATATGTTCAATC	2428
Db	174	-----	174
QY	2427	CCATGAGGAGGTGTTTCATCTTAGAACTCCCATGCAAGAGCTACATTAAAGAGAGCTGC	2368
Db	174	-----	174
QY	2367	AGGTTAAGGGCTTAGA-----GATGGAAACACAGGTGACTGAGTTATTTCAGCTCCCA	2314
Db	175	-----AspGlyAlaArgGlySerAspGlyGlnProGly	187
QY	2313	AAACCCCTTCTTAGTGTTCTCAACTAGGAGGCTAGCTGTTAACCCCTGAGCCTGGSTA	2254
Db	187	roGlyProPro--GlyThrAlaGlyPhePro	198
QY	2253	ATCCACTGTCAGAGTCCCGCATTCAGTGTCATGGAGCCCT-----TCTGGCCTCCCTG	2200
Db	198	erProGlyAlaLysGlyGluValGlyProAlaGlySerProGlySerAsnGlyAlaProG	218
QY	2199	TATAAGTCCAGACTGAACCCCTTGGGAAGGCTCCAGTCAGGCAGCCCTAGAGACTGGG	2140
Db	218	ly-----GlnArgGlyGluProGlyProGlnG	227
QY	2139	GAGAG-----AGGAGAGGGAGCCCGCCAGCCCGCTGCAGCTACGCACTCAGCAG	2086
Db	227	lyHisAlaGlyAlaGlnGlyProProGlyProProGlyIle-----AsnGlyS	243
QY	2085	CACAGGTGGCAGCAGGACCAATTTACTTTTGGCAGCAACAGAACTGGGGCCAGCCC	2026
Db	22	-----	22

114 ---GlyGly-----SerArgGlyAspProGlyProGlyAlaGlnGly 127  
128 ProAla---GlyProGlyGlySerArgGlyAspPro---GlyProPro----- 141  
2188 ACTGAAACCCCTTGGAGGCTCCAGTCAGCCAGGCTAGAGACTGGGAGAGAGAGA 2129  
142 -----GlyAlaGln-----GlyProAlaGlyProGlyGlySerArg 153  
2128 GGGAGCCCGCCAGCCCGCTGTCAGCTACGCTACGCTACGAGCAGCAGGCTGGCAGCAGA 2069  
154 GlyAspProGlyProGlyAlaGln-----GlyProAlaGlyProGlyGlySerArg 171  
2068 ---GAGCCACATTTACTTTGGCAGCAGCAGAACTGGCGGCGCCAGCCGCGCAGC----- 2020  
172 GlyAspProGlyProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAs 191  
2019 -----CCCATGGGCTAACAGGA---GCGGGAGCTGGGACCCAGTGGAGGAGGCCCT 1970  
191 pProGlyProGlyAlaGlnGlyProAla---GlyProGlyGlySerArgGlyAspPro 211  
1969 CCACCCCAATGTGCTGGAAGTTTCTACGCTGAGTATTTGGCCAACTGCTTCTGCTCAA 1910  
211 lyProGlyAlaGlnGly----- 217  
1909 TACTACCTGTGTAGCAAAATAATGGCGCAGCAGCCAGGCTGGCGCAGACACCATATA 1850  
218 -----ProAlaGlyProGlyGlySerArgGlyAspProGly-----ProProGlyA 233  
1849 GGCAGTACAGACTGGCTGAGCTGGCAATGGAGCCCATAAACAGGGATGGGCGCAGCTG 1790  
233 laGln-----GlyProAlaGlyProGlyGlySerArgGlyAspProGlyProG 250  
1789 GCACAGCAGGAGGCACTATCCAGCATGGCGA-----GGTCCAGGCAGATGCC 1742  
250 lyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyProGlyAlaH 270  
1741 CCGGCGCCGAA-----CCACCCTGGCGCTGGCTGAGCTCAGCCAC 1703  
270 isGlyProAlaGlyProGlyGlyAlaHisGlyProAlaGlyProGlyGlyAlaHisGly 290  
1702 CACACACCTAGCAGATCATCAG-----CCACTGCTCCAGCAGCCAGCTGCTCATTTAG 1608  
290 roAlaGlyProGlyGlyAlaHisGlyProAlaGlyProGlyGlyAlaGlnGlyPro----- 308  
1661 GCGCGGTGGAGTGGGAGCAGG-----CCACTGCTCCAGCAGCCAGCTGCTCATTTAG 1608  
309 --AlaGlyProGlyGlySerArgGlyAspProGlyProGlyAlaGlnGlyProAlaG 328  
1607 GG-----AAGGAGCTCCAGG-----TTAGGCGCTG 1581  
328 lyProGlyGlySerArgGlyAspProGlyProGlyAlaGlnGlyProGlyProGlyPro 348  
1580 GCAGGAGCTGCTCATCAGCTGCTCTACTGCTAGCAGCTCCAGCTGCTCCCTCGGTATT 1521  
348 ly-Gly-----SerArgGlyAspProGlyProGlyProGlyAlaGlnGlyProAlaGlyPro 365  
1520 TGGGCA-----GGAAACACTGCTTCTCCCGTGGTAGAGGAGGCCAGTGTGTAGGCA 1467  
366 GlyGlySerArgGlyAspProGlyProGlyAlaGlnGly-----ProAla 381  
1466 GGAATGTCAGGCTGAGAGGTGAACCCGGTGGAGGCGGCTGAGCTGTCTACCCAGCGCA 1407  
382 GlyProGlyGlySerArg-GlyAspProGlyProGly----- 394  
1406 CACTGTGGACAGGATGTGGCAGCGCAGCAGCCAGGGAAGCTGCCACACTGGCAAT 1347  
395 --AlaGlnGlyProAla---GlyProGlyGlySerArgGlyAsp---ProGlyProGly 412  
1346 AGACTGCTGAGTGGCGAATCGCTGCACCCAGCGCTCCATGACCCAGAGAGA----- 1296  
413 -----AlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspPro 426

1295 ---AGACCAGGAGATGGCGCACTGCAGGAACAGCC-----CCAGGCTGC 1254  
427 GlyProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyPro 446  
1253 CCATCCGAACGCCCTTCATCATAGTGTCTCCGGGCTCGGTGCGCGCTCAGCTCGGGCA 1194  
447 Pro-----GlyAlaGlnGlyProAlaGlyProGlyGlySerArgGly 460  
1193 CGCCCTGGTACAGCCCTCGCCCAACCAATCCGTGTAAACAGCGTGAAGGTATCATGATG 1134  
461 AspProGly---ProProGlyAlaGlnGlyProAlaGly----- 472  
1133 CCATCCAGTGCAGCTCAGCCAGGAGCGCGCAGGCTGGCGGCTATCGGCAGC 1074  
473 ProGlyGlySerArgGlyAspProGlyProGlyAlaGln---GlyProAlaGly--- 490  
1073 ACAGCTGGTGCAGCCCGGAGAGCGGCGCCAGTTCCGGAAAGCAAGCGGCGCCGCGC 1014  
491 ProGlyGlySerArgGlyAspProGlyProGlyAlaGlnGlyProAlaGlyProGly 510  
1013 -----ATGGACAGCAGTGGCGGCGACAAGAGGGGCGCCAGACCCCTTCTGCTGCT 963  
511 GlySerArgGlyAspProGlyProGlyAlaHisGlyProAlaGlyProGlyGlyAla 530  
962 CGTGGGCGCCAGCGTGCCT-----CCTCAGCCAGCAGCAGTGTGCTG 918  
531 His---GlyProAlaGlyProGlyGlyAlaHisGlyProAlaGlyProGly----- 546  
917 CTAGCAGGCTGAGAGATGAGGTGAGCAGGAGGCGCAAGAGGCGCACT----- 873  
547 -----GlyAlaHisGlyProAlaGlyProGlyGlyAlaGlnGlyProAlaGly 562  
872 -----CCTCTGGTGGTCCAGGT-----AGGGGCGCCAGG 843  
563 ProGlyGlySerArgGlyAspProGlyProGlyAlaGlnGlyProAlaGlyPro----- 581  
842 CACTGGTGTCCAGTCAATGGCAGGAGGAGGTAGCCAGGAGGCGCCAGCAGCTGATCA 783  
582 -----GlyGlySerArgGlyAspPro----- 588  
782 TGAAGCATACAGACTAGGCTGGCGCACAGTGTCCG---GCTCCCGGAAGAGGTGCT 726  
589 -----GlyProGlyAlaGlnGlyProAlaGlyProGly---GlySer 602  
725 AGCAGGCGCTCCAGTGGAGTGAAGCACACCTGGCCACAGAGTCCAGCAGGCGCCA--- 669  
603 ArgGlyAspPro-----GlyProGlyAlaGlnGlyProAlaGlyProGly 618  
668 -----CGCCAGGATGAGCAGTGCAGCT---CCAGGCGCC--- 636  
619 GlySerArgGlyAspProGlyProGlyAlaGlnGlyProAlaGlyProGlyGlySer 638  
635 ---TGGATCCGGCAGCAGCAGCCCTGCTAGCCAGCGCGCTTGGATGAGAAGAGG 579  
639 ArgGlyAspProGlyProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGly 658  
578 TCACAGGATCCCAAGCAGACAGTCCAGATGAGGCGCGCGGCGCCATAGGTCCAC 519  
659 AspProGlyPro-----GlyAlaGlnGlyProAla 669  
518 GCCAGTGTACTGGCTGAGCCTAGAGCGGCGCACACAGCAGCCCGCCAGCAGCTGACCAA 459  
670 GlyProGly-----GlySerArgGlyAspProGlyProGlyProGlyAlaGln 684  
458 TGGCAGCAGCAGTGTATGAACTTCTCTACCCCGCTTCCAGCAGCAGGAGCGGCA 399  
685 GlyProAlaGlyProGlyGlySerArgGlyAspProGlyProGlyAlaGlnGly--- 703  
398 CATAGGTGATCCCTGGCGGCGCAACACACCTCCAGGCGCAAGGTTAGCAGGTTGACCA 339  
704 -----ProAlaGlyProGlyGlySerArg-----GlyAspProGly 715

```

360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378
1178 CTTGCCCGCCAGAAATCGTGTAAACAGACGCTGAAGTCAATGAGTCCATCCAGC-----1125
379 GlyAlaProGlyThrProGlyProGlnGly-----LeuProGlySerProGly 394
1124 -----TGCACAGCTCAGCCAGCAAGAGCGCGCAGGCTGGCGGCGCAGCAGCA 1071
395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414
1070 GCT-----GGTCAGCGCGGGAAGCAGAGCGCGCCAGGTTCCGGAAGCAGCGGG---1020
415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434
1019 CCCGGCATGACAGCAGTGGGGCGACAGAGAGGGGGCGCAGCCCTTCTGCTGGCTCGG 960
435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453
959 TGGGGCCCGCAGCTGCTCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGAGGTGAGCA 903
454 GlyAlaProGlyThrProGlyProGln-----Gly 463
902 AGATAGGGTGAAGAGCCAAAGAGGACACTCT---CCTGGGTGCCCA 858
464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483
857 GGTAGGGGGCAGGGCACTGCTCCAGTCAATGCGCAGGAGGTAGCCAGGACG 798
484 Gly---AlaProGly-----ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500
797 CCCCAGACTGATCATGAAGGCATAGACAGAGTAGGCTGGCGACAGTGT---CCGGGT 741
501 Pro-----GlyThrProGlyProGlnGlyLeuProGly 511
740 ---CCCGGAAGAGTCAAGAGCAGGCGCTCCAGTGGAGTGAAGCACACTGGCCACAGA 684
512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly-----529
683 AGTCCAGCAGCCCGCAGGAGTGAAGCAGTCCAGTCCAGGCGCGCTGGGATCCGGCG 624
530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla-----ProGly 547
623 ACAGCAGCCCTGTAGCAGCCCGCCTTGGGATGAGAAAGAGGCTCAGCAGGATCCCA 564
548 ThrProGly-----ProGlnGly---LeuProGlySerProGlyAlaPro 561
563 AGACACTG---CCGACATGAAGGCGCGCGCGCCCATAGCTCCACGCCAGTGTGTAC 507
562 GlyThrProGlyProGlnGlyLeuProGly-----571
506 TGCTGAGCCTAGGAGCGGCGACACAGCAGGCGCCAGCAGTGGCAATCCCGCAGCA 447
572 -----SerProGlyAlaProGlyThrProGlyProGln-----GlyLeuPro 585
446 TGCTCATGAATCTCTCTACCCCA-----CTTCCAGCAGCAGAGCGCGCA 399
586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605
398 CATAGGTGATCGCTGGCGGCAACACACTCCAGGCGCAAGGTTAGAGTTGACAGCA 339
606 -----ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619
338 AGACTGGCTTCCGGT---GCCGACAGGCGGCTCACCCACAGCCCTCTGGACCATAG 282
620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639
281 TGGGCCAGGCGGG---TAGGGCTCAGGGGCGCTTCAGGCACATCCAGAACTGCTGCT 225
639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659
224 CGGCTCTGCTCCAGAACTCGGCGCTCTCTCTCTGCTGCGCCAACTGCCTAGGAATCA 165

```

659 uProGlySerProGlyAlaProGlyThrProGly-----ProGlnGlyLe 674  
164 GCCAGCGCCCATTTCTCCAGCCCTTTGGTGGCGGTCCAGCTTCTCAGCCCATGCTCA- 106  
674 uProGlySerProGlyAla---ProGlyThrProGlnGly-LeuProGlySerP 693  
105 -----ACACCTGCTGCTGGGGCACCCTCAGTGGGGACACAGTCTCATCACT 60  
693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712  
59 CAGATCTCTGGC---CGAGGCGCGGCTGTCCACCCGGAGCC 22  
712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725

RESULT 41  
US-08-642-255-72  
; Sequence 72, Application US/08642255  
; Patent No. 5773249  
; GENERAL INFORMATION:  
; APPLICANT: CAPPELLO, Joseph  
; APPLICANT: FERRARI, Franco A.  
; TITLE OF INVENTION: High Molecular Weight Collagen-Like  
; TITLE OF INVENTION: Protein Polymers  
; NUMBER OF SEQUENCES: 135  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/642,255  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROWLAND, Bertram I.  
; REGISTRATION NUMBER: 20,015  
; REFERENCE/DOCKET NUMBER: A5556-3/BIR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 494-8700  
; TELEFAX: (415) 494-8771  
; INFORMATION FOR SEQ ID NO: 72:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1065 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-642-255-72

Alignment Scores:  
Pred. No.: 4,82e-19 Length: 1065  
Score: 353.50 Matches: 270  
Percent Similarity: 32.33% Conservative: 20  
Best Local Similarity: 30.10% Mismatches: 332  
Query Match: 5.68% Indels: 276  
DB: 1 Gaps: 61

US-09-759-143-110 (1-3410) x US-08-642-255-72 (1-1065)  
QY 2347 GSGAAACAGCTGACTGAGTTTATTCAGCTCCCAAAACCCCTTCTTAGTGTGCTCTCAA 2288  
Db 100 GlyAspProGlyProGlyAlaGlnGlyProAlaGlyPro-----113  
QY 2287 CTAGGAGGCTAGCTGTTAACCCCTGAGCCTGGGTATCA-----2249



Db 74 LeuProGlySerPro---GlyAlaPro-----GlyThrProGlyPro---GlnGly 88  
QY 2167 CTCCAGTCAGGAGCCCTAGAGACTGGGAGAGAGAGGAGCGCCAGCCAGCT 2108  
Db 89 Leu---ProGlySerPro-----GlyAlaProGlyThrProGly 100  
QY 2107 GTGCAGCTAGCAGCTCAGCAGACAGGAGGTGGCAGCAGAGCCATTTGGCAGC 2048  
Db 101 ProGln-----GlyLeuProGlySerProGlyAlaProGlyThrProGly--p 116  
QY 2047 AACAGAACTGGCGCCAGCCGCGCAGCCCATGGGCTAACAGGAGCGGGAGCTGGGA 1988  
Db 116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136  
QY 1987 CCCAGTGAGGAGCCCTCCACCCCAATGTGCTGGAAGTTTCTACGCTCAGTATTGGC 1928  
Db 136 lySerProGlyAlaProGlyThrPro----- 144  
QY 1927 CAAGTCGCTTGTCAATACTACCTGTGTAGCAAGTAATGGCCAGCAGCCAGGCC 1868  
Db 145 -----GlyProGlnGlyLeuPro-----GlySerProGlyAla 155  
QY 1867 TGGCGGACACCATATAGCAGCTGACAGACTGGCTGAGCTGGCAATGAGGCCATAAA 1808  
Db 155 laProGlyThrProGlyProGlnGlyLeuProGly-----SerPro----- 168  
QY 1807 CAGGGATGGGCGCACCTGGGACAGCAGGAAGGCAC-----TATC 1769  
Db 169 -----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 186  
QY 1768 CAGGATGGCGAGCTCCAGGAGATGCCCGGC-----CCGGAACACCCCTGGC 1721  
Db 186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 206  
QY 1720 CTCGGTG-----GGCTACCCACCACACAGCTACGGAGACATCACAGGCA-----G 1674  
Db 206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226  
QY 1673 AGCCCGGAGAGCGGGTGGAGGTGGGAGAGCCATGCTCCAGCAGCCAGCTGTC 1614  
Db 226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 246  
QY 1613 CATTAGGAAAGGAGCTCCAGGCTTAGGCTGGCAGGAAGCTGGTCATCAGCTGCTCT 1554  
Db 246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264  
QY 1553 CACTGTAGCACCTCCAGTGTCCCTCGGTATT-----TGGGCGAAGAACCTGCTCT 1500  
Db 265 Gly-----ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278  
QY 1499 CCCGGTGGTAGAGGAGCCAGTGTAGGGCAGGATCTGAGGGCTGAGAGGTGAACC 1440  
Db 279 ProGlyProGlnGlyLeuPro-----GlySerProGlyAla-ProGlyThrPr 294  
QY 1439 CGGTGAGGGCGGTGAAGTGTCCACCAGGCCACACTGTGGGACAGGCATGCGCAC-- 1382  
Db 294 oGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPr 309  
QY 1381 -GSCAGCCAGGAAAGCTGCCACACTGGCCCAATAGACTGCTCAGTGCAGTCCGAATCGCT 1323  
Db 309 oGly--ProGlnGly--LeuProGlySerPro----- 319  
QY 1322 GCACCGCGGTCCATGACAGAGAGAACAGCAGGAGATGGCGCACTGCAGAACACA--- 1266  
Db 320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339  
QY 1265 -----GCCCGAGCTGCCATCC-----GAACGCCTT 1239  
Db 340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359  
QY 1238 CATCATAGTGTCTCCGGGCTCGGTGCCCGGCTCAGCTCTGGGCGACCGCTCGGTACAGCC 1179  
Db 74 LeuProGlySerPro---GlyAlaPro-----GlyThrProGlyPro---GlnGly 88  
QY 2167 CTCCAGTCAGGAGCCCTAGAGACTGGGAGAGAGAGGAGCGCCAGCCAGCT 2108  
Db 89 Leu---ProGlySerPro-----GlyAlaProGlyThrProGly 100  
QY 2107 GTGCAGCTAGCAGCTCAGCAGACAGGAGGTGGCAGCAGAGCCATTTGGCAGC 2048  
Db 101 ProGln-----GlyLeuProGlySerProGlyAlaProGlyThrProGly--p 116  
QY 2047 AACAGAACTGGCGCCAGCCGCGCAGCCCATGGGCTAACAGGAGCGGGAGCTGGGA 1988  
Db 116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136  
QY 1987 CCCAGTGAGGAGCCCTCCACCCCAATGTGCTGGAAGTTTCTACGCTCAGTATTGGC 1928  
Db 136 lySerProGlyAlaProGlyThrPro----- 144  
QY 1927 CAAGTCGCTTGTCAATACTACCTGTGTAGCAAGTAATGGCCAGCAGCCAGGCC 1868  
Db 145 -----GlyProGlnGlyLeuPro-----GlySerProGlyAla 155  
QY 1867 TGGCGGACACCATATAGCAGCTGACAGACTGGCTGAGCTGGCAATGAGGCCATAAA 1808  
Db 155 laProGlyThrProGlyProGlnGlyLeuProGly-----SerPro----- 168  
QY 1807 CAGGGATGGGCGCACCTGGGACAGCAGGAAGGCAC-----TATC 1769  
Db 169 -----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 186  
QY 1768 CAGGATGGCGAGCTCCAGGAGATGCCCGGC-----CCGGAACACCCCTGGC 1721  
Db 186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 206  
QY 1720 CTCGGTG-----GGCTACCCACCACACAGCTACGGAGACATCACAGGCA-----G 1674  
Db 206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226  
QY 1673 AGCCCGGAGAGCGGGTGGAGGTGGGAGAGCCATGCTCCAGCAGCCAGCTGTC 1614  
Db 226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 246  
QY 1613 CATTAGGAAAGGAGCTCCAGGCTTAGGCTGGCAGGAAGCTGGTCATCAGCTGCTCT 1554  
Db 246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264  
QY 1553 CACTGTAGCACCTCCAGTGTCCCTCGGTATT-----TGGGCGAAGAACCTGCTCT 1500  
Db 265 Gly-----ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278  
QY 1499 CCCGGTGGTAGAGGAGCCAGTGTAGGGCAGGATCTGAGGGCTGAGAGGTGAACC 1440  
Db 279 ProGlyProGlnGlyLeuPro-----GlySerProGlyAla-ProGlyThrPr 294  
QY 1439 CGGTGAGGGCGGTGAAGTGTCCACCAGGCCACACTGTGGGACAGGCATGCGCAC-- 1382  
Db 294 oGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPr 309  
QY 1381 -GSCAGCCAGGAAAGCTGCCACACTGGCCCAATAGACTGCTCAGTGCAGTCCGAATCGCT 1323  
Db 309 oGly--ProGlnGly--LeuProGlySerPro----- 319  
QY 1322 GCACCGCGGTCCATGACAGAGAGAACAGCAGGAGATGGCGCACTGCAGAACACA--- 1266  
Db 320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339  
QY 1265 -----GCCCGAGCTGCCATCC-----GAACGCCTT 1239  
Db 340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359  
QY 1238 CATCATAGTGTCTCCGGGCTCGGTGCCCGGCTCAGCTCTGGGCGACCGCTCGGTACAGCC 1179  
Db 360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378  
QY 1178 CCTCCCGCCAGCAAAATCCCTGTAAACACACGCTGAAGCTCATGAGTCCCATCCAGC----- 1125  
Db 379 GlyAlaProGlyThrProGlyProGlnGly-----LeuProGlySerProGly 394  
QY 1124 -----TGCACAGCTCAGCCACAGAGAGCCGCGCAGGCTGGGGGATCGCCAGCACA 1071  
Db 395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414  
QY 1070 GCT-----GGTGCAGCCGGGAGCAGGCGCCAGCTCCGGAAGCCAAAGCGGG--- 1020  
Db 415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434  
QY 1019 CCCGCGATGAGCAGAGTGGGCGCAGAGAGGGGGGCGCAGACCCCTTCTCTGCTGCTGG 960  
Db 435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453  
QY 959 TGGGCGCCAGCGCTGCCT---CCTCAGCCACACAGCAGTGTGGCTGTACGCAAGTGAGGA 903  
Db 454 GlyAlaProGlyThrProGlyProGln-----Gly 463  
QY 902 AGATGAGGCTGAGCAGCGCCAAAGAGGCACTCCT-----CCTGGGTGCCCA 858  
Db 464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483  
QY 857 GGTAGGGGCGCCAGGCACTGGTGTCCAGTCAATGGCAGGAGGAGGTAGCCAGGCGAGC 798  
Db 484 Gly---AlaProGly-----ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500  
QY 797 CCCCAAGACTGATCATGATGAAGCATAGACAGAGTAGGCTTGGCGCAGCAGTGGT---CCGGGT 741  
Db 501 Pro-----GlyThrProGlyProGlnGlyLeuProGly 511  
QY 740 ---CCCGAAGAGGTCAGAGAGCGGCTCCCTGAGTGAAGCAGCAGCTGGCCACAGA 684  
Db 512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly----- 529  
QY 683 AGTCAGCAGCCCGCCAGCGCCAGGATGAGCAGTCCAGCTCCAGGGGCTTGGATCCGGC 624  
Db 530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla-----ProGly 547  
QY 623 ACAGCAGCCCTGCTAGCAGCGCGGCTTGGATGAGAAAGAGCTCAGCAGGATGCCCA 564  
Db 548 ThrProGly-----ProGlnGly---LeuProGlySerProGlyAlaPro 561  
QY 563 AGGACAGTG---CCACAGATGAAGGCGCGCGCGCCATAGCTCCAGCGGCTGGTCA 507  
Db 562 GlyThrProGlyProGlnGlyLeuProGly----- 571  
QY 506 TGGCTGAGCCCTAGGAGCGGGACACAGACCGCCAGCAGCTGGACCAATGCCAGCACCA 447  
Db 572 ---SerProGlyAlaProGlyThrProGlyProGln-----GlyLeuPro 585  
QY 446 TGGTCATGAATCTCTCTCTACCCCA-----CTTCCAGCAGCAGAGCGCGCA 399  
Db 586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605  
QY 398 CATAGTGATGCTCGCGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACACCA 339  
Db 606 -----ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619  
QY 338 AGACTGGGCTTTCCGCT---GCCGAGCAGCGGCTCACCCAGCAGCTCTGGACCATAG 282  
Db 620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639  
QY 281 TGGGCGGCGGG---TAGGCTCAGGGCGGCTTACAGGCACTCCCACTGCTTCTCTCT 225  
Db 639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659  
QY 224 CGGCTCTGCTCCAGAACTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 165  
Db 659 uProGlySerProGlyAlaProGlyThrProGly-----ProGlnGlyLe 674





MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,085B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/927,258  
FILING DATE: 04-NOV-1986  
PRIOR APPLICATION DATA: US 07/114,618  
FILING DATE: 29-OCT-1987  
APPLICATION NUMBER: US 08/053,049  
FILING DATE: 22-APR-1993  
APPLICATION NUMBER: US 08/175,155  
FILING DATE: 29-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-55186-6/RFT/MTK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 837 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-08-482-085B-103

Alignment Scores:  
Pred. No.: 2,61e-19 Length: 837  
Score: 356.50 Matches: 263  
Percent Similarity: 35.16% Conservative: 32  
Best Local Similarity: 31.35% Mismatches: 327  
Query Match: 5.73% Indels: 218  
DB: 3 Gaps: 57

US-09-759-143-110 (1-3410) x US-08-482-085B-103 (1-837)

QY 2347 GGAAACAGGTGACTGAGTTATTACCTGAGTCCCAAAACCTTCTCTAGTGTGCTCAA 2288  
DB 46 GlySerProGly-----AlaProGlyThrPro-----GlyProGly 58  
QY 2287 CTAGGAGCTAGCTGTTAACCTGAGCTGGGTATCCACCTGCAGAGTCCCGCATGCC 2228  
DB 59 LeuProGly-----SerProGlyAlaProGlyThrProGlyProGly 73  
QY 2227 AGTGCATGAGCCCTTCTGCTCCCTGATAGTCCAGACTGAAACCCCTTGAAGGC 2168  
DB 74 LeuProGlySerPro--GlyAlaPro-----GlyThrProGlyPro--GlyGly 88  
QY 2167 CTCCAGTCAGGACCCCTAGACTGGGAGAGAGGAGGAGCCGCCAGCCCGCCAGCT 2108  
DB 89 Leu---ProGlySerPro-----GlyAlaProGlyThrProGly 100  
QY 2107 GTCCAGCTAGCCACTCAGCAGCACAGGAGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048  
DB 101 ProGln-----GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116  
QY 2047 AACAGAACTGGCGGACGCGGAGCCCGCTTGGGCTTAACAGGAGGCGGAGCTGGGA 1988  
DB 116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136  
QY 1987 CCAGTGAGGAGCCCTCCACCAATGTCTGGAAGTTTCTACGCTGAGTATTGGC 1928  
DB 136 lySerProGlyAlaProGlyThrPro-----144

QY 1927 CAAGTCGCTCTTGTCAAAATACTACTGTGTAGCAAAATAGTAAATGCGCAGACCCAGGCC 1868  
DB 145 -----GlyProGlnGlyLeuPro-----GlySerProGlyAla 155  
QY 1867 TCGCGCAGACACCATATAGGCGAGTACAGACTGGCTGAGTGGACATGAGGCCCATAAA 1808  
DB 155 laProGlyThrProGlyProGlnGlyLeuProGly-----SerPro---- 168  
QY 1807 CAGGATGGCGCCACCTGGGACACAGCAGGAAGGCAC-----TATC 1769  
DB 169 -----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 186  
QY 1768 CAGGATGGCGAGGTCCAGGCAGATGCCCGGC-----CCGGAACACCTGGC 1721  
DB 186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 206  
QY 1720 CTCGGTG-----GGTCACCCACCCACACACACAGTACGAGACATCACAGCA-----G 1674  
DB 206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226  
QY 1673 AGGCCCGCAGAGCGGGTGGAGTGGAGCAGGCACCTGCCTCCAGCAGCCACGTC 1614  
DB 226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 246  
QY 1613 CATTAGGAAGGAGCTCCAGGCTTGGGCTGGCAGGAAGCTGGTGCATCAGGCTGTCT 1554  
DB 246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264  
QY 1553 CACTGCTAGCACCTCCAGTGTCCCTCGGTATT-----TGGCAGGAACACCTGCTTCT 1500  
DB 265 Gly-----ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278  
QY 1499 CCCGTTGTAGAGGAGGCCAGTGTGTAGGCGAGATCTGCAGGCTGAGAGGTGAACC 1440  
DB 279 ProGlyProGlnGlyLeuPro-----GlySerProGlyAla-ProGlyThrPr 294  
QY 1439 CGGTAGGCGGCTGAAGCTGTCCACACGCGGCACACTGTGGGACAGACTGTGGCACC-- 1382  
DB 294 oGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPr 309  
QY 1381 -GGCAGCCACAGGGAAGCTCCACACTGGCCCAATAGACTGCTCGAGTCCCGATCGCT 1323  
DB 309 oGly--ProGlnGly---LeuProGlySerPro-----Gly 319  
QY 1322 GCACGAGCGGTCCATCAGCAGAGAGAGAGAGGAGATGGCGACTGCAGGAACA--- 1266  
DB 320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339  
QY 1265 -----GCCCCAGGTGCCCATCC-----GAACGCTT 1239  
DB 340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359  
QY 1238 CATCATAGTGTCTCGGCGCTCGGTCCGCTCAGCTCGTGGGACCCCTCTGTGTACAGCC 1179  
DB 360 ProGlySerProGlyAlaProGlyThrProGlyPro--GlnGlyLeuProGlySerPro 378  
QY 1178 CCTCCGCCACGAATCCGTGTAAACAGCCTGAAGTGCATGAGTGCATCAGTGCATCAGCAGCAGC 1125  
DB 379 GlyAlaProGlyThrProGlyProGlnGly-----LeuProGlySerProGly 394  
QY 1124 -----TGCACAGCTCAGCCACAGAGAGCGCGCCAGGCTGGGGGATCGCGCAGCACA 1071  
DB 395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414  
QY 1070 GCT-----GCTGCAGCGCGGAGAGCGCGCCAGTTCCTCCGAAAGCCAGCGGG-- 1020  
DB 415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434  
QY 1019 CCCGCGCATGGACAGCTGGGGGCGCACAGGAGGGGCGCCAGAGCCCTTCTGCTGGCTCGG 960  
DB 435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453



QY 398 CATAGTGATGCTGGCGCCAAACACACACTCCAGCCCAAGGTTAGCAGGTTGACGACA 339  
 Db 606 -----ProGlyThrProGlyProGlyLeuProGlySerProGly 619  
 QY 338 AGAGCTGGCTTTCGGT---GCCGACAGCGCGCTCACCCACAGCCCTCTGGACCATAG 282  
 Db 620 AlaProGlyThrProGlyProGlyLeuProGlySerProGlyAlaProGlyThr-pr 639  
 QY 281 TGGGCGAGCGGG---TAGGGCTCAGGGCGGTTTCAGGCACTCCAGAACTGCTCGGPT 225  
 Db 639 oGlyProGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlyLe 659  
 QY 224 CGGCTCTCTCAGAGCTGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 165  
 Db 659 uProGlySerProGlyAlaProGlyThrProGly-----ProGlyLe 674  
 QY 164 GCCAGCGCCATTTCTGCCAGCCCTTTGGTGGCGGCTCCAGCTTCTCAGCCCATGCTCA- 106  
 Db 674 uProGlySerProGlyAla---ProGlyThrProGlyProGlyLeuProGlySerP 693  
 QY 105 -----ACACTGCTGCTGTGGGCACTCAGTGGGACACAGCTCTCATCACT 60  
 Db 693 roGlyAlaProGlyThrProGlyProGlyLeuProGlySerProGlyAlaProG 712  
 QY 59 CAGATCTCTGGC---CGAGCGCGCGGCTGTCACCGGAGCC 22  
 Db 712 lyThrProGlyProGlyLeuProGlySerProGlyAla 725

## RESULT 37

US-08-707-237A-75  
 ; Sequence 75, Application US/08707237A  
 ; Patent No. 5830713  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ferrari, Franco A.  
 ; APPLICANT: Capello, Joseph  
 ; APPLICANT: Crissman, John W.  
 ; APPLICANT: Dorman, Mary A.  
 ; TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC  
 ; TITLE OF INVENTION: REPETITIVE DNA  
 ; NUMBER OF SEQUENCES: 108  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fehr, Hohnbach, Test, Albritton & Herbert  
 ; STREET: Four Embarcadero Center, Suite 3400  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States  
 ; ZIP: 94111-4187  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/707,237A  
 ; FILING DATE: 03-SEP-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/175,155  
 ; FILING DATE: 29-DEC-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/053,049  
 ; FILING DATE: 22-APR-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/609,716  
 ; FILING DATE: 06-NOV-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/269,429  
 ; FILING DATE: 09-NOV-1988  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/114,618  
 ; FILING DATE: 29-OCT-1987  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 06/927,258

US-09-759-143-110 (1-3410) x US-08-707-237A-75 (1-837)

QY 2347 GCGAAACAGGTCAGTGTATTTCAGCTCCCAAAACCTTCTCTAGGTGTGTCAA 2288  
 Db 46 GlySerProGly-----AlaProGlyThrPro-----GlyProGly 58  
 QY 2287 CTAGGAGGCTAGCTGTTAACCCCTGAGCCTGGGTAAATCCAGAGTCCCGCATTC 2228  
 Db 59 LeuProGly-----SerProGlyAlaProGlyThrProGlyProGly 73  
 QY 2227 AGTGCATGGAGCCCTTCTGGCCTCCTGTATAGTCCAGACTGAACCCCTTGAAGGC 2168  
 Db 74 LeuProGlySerPro---GlyAlaPro-----GlyThrProGlyPro---GlyGly 88  
 QY 2167 CTCAGTCAGGAGCCCTAGAGACTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2108  
 Db 89 Leu---ProGlySerPro-----GlyAlaProGlyThrProGly 100  
 QY 2107 GTGCAGCTACGCACTCAGCAGCAGAGGCTGGCAGAGAGAGAGAGAGAGAGAGAGAG 2048  
 Db 101 ProGln-----GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116  
 QY 2047 AACAGAACTGGCGCCAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1988  
 Db 116 roGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlyProGlyLeuProG 136  
 QY 1987 CCCAGTGGAGGAGCCCTCCACCCCAATGTGTGGAAGTTTCTACGCTGAGTATTGGC 1928  
 Db 136 lySerProGlyAlaProGlyThrPro-----GlyAlaProGlyThrPro----- 144  
 QY 1927 CAAGTCGCTTGTCAATACCTACCTGTGTAGCAAGTAATGGCGACAGAGAGAGAGAG 1868  
 Db 145 -----GlyProGlyLeuPro-----GlySerProGly 155  
 QY 1867 TGGCGAGACACCATATAGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1808  
 Db 155 laProGlyThrProGlyProGlyLeuProGly-----SerPro----- 168  
 QY 1807 CAGGATGGGCACTCTGGAG 1769  
 Db 169 -----GlyAlaProGlyThrProGlyProGlyLeuProGlySerProGlyAlaP 186  
 QY 1768 CAGGATGGGAGTCCAGGAG 1721  
 Db 186 roGlyThrProGlyProGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206  
 QY 1720 CTCGGTG-----GGCTCAGCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1674

Alignment Scores: 2.61e-19 Length: 837  
 Pred. No.: 356.50 Matches: 263  
 Score: 35.16% Conservative: 32  
 Percent Similarity: 31.35% Mismatches: 327  
 Best Local Similarity: 5.73% Indels: 218  
 Query Match: 2 Gaps: 57  
 DB: 57

US-08-707-237A-75  
 ; FILING DATE: 04-NOV-1986  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Treccartin, Richard F.  
 ; REGISTRATION NUMBER: 31,801  
 ; REFERENCE/DOCKET NUMBER: A-55186-10/WHO  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 781-1989  
 ; TELEFAX: (415) 398-3249  
 ; TELEX: 910 277299  
 ; INFORMATION FOR SEQ ID NO: 75:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 837 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-08-707-237A-75

Best Local Similarity:	31.35%	Mismatches:	327
Query Match:	5.73%	Indels:	218
DB:	1	Gaps:	57
US-09-759-143-110 (1-3410) x US-08-642-255-101 (1-837)			
QY	2347	GGGAAACAGGCTGACTGAGTTATTCAGCTCCCAAAACCTTCTCTAGGTGTCTCAA	2288
DB	46	GlySerProGly-----AlaProGlyThrPro-----GlyProGlnGly	58
QY	2287	CTAGGAGGCTAGCTGTTAAACCTGAGCTGGGTAAATCCACCTGCAGAGTCCCGCATTC	2228
DB	59	LeuProGly-----SerProGlyAlaProGlyThrProGlyProGlnGly	73
QY	2227	AGTCATGGAGCCCTTCTGGCTTCTGTTATAGTCCAGACTGAAACCCCTTGGAGGC	2168
DB	74	LeuProGlySerPro-----GlyAlaPro-----GlyThrProGlyPro-----GlnGly	88
QY	2167	CTCCAGTCAGGACCCCTAGAGACTGGGGAGAGAGAGGAGCGGAGCCGCCACCT	2108
DB	89	Leu-----ProGlySerPro-----GlyAlaProGlyThrProGly	100
QY	2107	GTGACGTACACCTCAGCAGCAGAGGTGGCAGCAGAGCCACATTTCTGGCAGC	2048
DB	101	ProGln-----GlyLeuProGlySerProGlyAlaProGlyThrProGly--P	116
QY	2047	AACAGAACTGGGGCCAGCCGCGGAGCCCATGGGCTAACAGAGCGGGAGCTGGGA	1988
DB	116	roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG	136
QY	1987	CCCAGTAGGAGGCGCTCCACCCCAATGTCTGGAAGTTTCTACGCTGAGTATTGGC	1928
DB	136	lySerProGlyAlaProGlyThrPro-----	144
QY	1927	CAAGTCGCTCTGTCAAATACCTACTGTGTAGCAAAAGTAATGCGGACCCAGCC	1868
DB	145	-----GlyProGlnGlyLeuPro-----GlySerProGlyAla	155
QY	1867	TGCGCAGACACATATAGGAGCTGACAGACTGGCTGAGCTGGCAATGGAGCCATAAA	1808
DB	155	laProGlyThrProGlyProGlnGlyLeuProGly-----SerPro-----	168
QY	1807	CAGGATGGGCGCCTGGCAGCAGGAGGAGCAG-----TATC	1769
DB	169	-----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP	186
QY	1768	CAGGATGGGCGCCTGGCAGCAGGAGGAGCAG-----CCGGAACCCCTGGC	1721
DB	186	roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP	206
QY	1720	CTCGGTG-----GGCTACCCACCACACACGTCAGGAGACATCAGAGCA-----G	1674
DB	206	roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG	226
QY	1673	AGGCCCCAGAGCGGGGTGGAGTGGGAGCAGCCACTCCCTCCACACCCAGGTGC	1614
DB	226	lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP	246
QY	1613	CATTAGGGAAGGAGCTCCAGGCTTAGGGCTGGCAGGAGGCTGTCATCAGCTGTCT	1554
DB	246	roGlyThrProGlyProGlnGlyLeu-----ProGly-SerProGlyAlaProGlyThrP	264
QY	1553	CACGTGTAGCACCTCCAGTGTCCCTCGGTATT-----TGGGAGGAGACCTCTCTCT	1500
DB	265	Gly-----ProGlnGlyLeuProGlySerProGlyAlaProGlyThr	278
QY	1499	CCGGTGTAGAGGAGGAGCTGTAGGCGGAGTCTGAGGCTGAGAGGTGAACC	1440
DB	279	ProGlyProGlnGlyLeuPro-----GlySerProGlyAlaProGlyThrPr	294
QY	1439	CGGTGAGGCGGCTGAAGCTGTCCACCGCCGACACTGTGGGAGGAGCTGTGGCACC	1382
DB	294	oGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPr	309
QY	1381	-GGCAGCCACAGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTCCGCAATCGCT	1323
DB	309	oGly--ProGlnGly---LeuProGlySerPro-----Gly	319
QY	1322	GCACACGCGGTCCATACAGAGAGAACACAGGAGATGCGGACTGCAGAGACACA----	1266
DB	320	AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro	339
QY	1265	-----GCCAGGCTGCCCATCC-----GAACGCGCT	1239
DB	340	GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu	359
QY	1238	CATCATAGTGTCTCCGGCTCGGTGCCCGGTCTAGCTCTGGGACAGCCCTGTTACAGCC	1179
DB	360	ProGlySerProGlyAlaProGlyThrProGlyPro-----GlnGlyLeuProGlySerPro	378
QY	1178	CTCGGCCACGAATCCGTGTAAACAGCGTGAAGTCAATGAGTGCATCCACCC	1125
DB	379	GlyAlaProGlyThrProGlyProGlnGly-----LeuProGlySerProGly	394
QY	1124	-----TGACAGCTCAGCCACGAGAGCGGCGCAGGCTGGCGGATCGCGGACACACA	1071
DB	395	AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro	414
QY	1070	GCT-----GGTCAGCCGCGGGAAGCAGGCGGCCAGGTTCCGGAAGCAAGCGGG	1020
DB	415	GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu	434
QY	1019	CCCGGATGACAGCAGTGGGCGCAGAGGAGGCGGCGGACACCCCTCTGCTGGCTCGG	960
DB	435	ProGlySerProGlyAlaProGlyThr-----ProGlyProGlnGlyLeuProGlySerPro	453
QY	959	TGGGCGCCACGCTGCTCT---CCTCAGCCACACAGCAGTGTGTGCTGCTACGAGTGAGGA	903
DB	454	GlyAlaProGlyThrProGlyProGln-----Gly	463
QY	902	AGTAGGCTGACAGGCGCCAAAGAGCAGCTCT---CCTGGGTGCCCA	858
DB	464	LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro	483
QY	857	GGTAGGCGGCGCAGGCGCCTGCTCCAGTCAATGGCAGGACGAGGATAGCCAGGCGC	798
DB	484	Gly---AlaProGly-----ThrProGlyProGlnGlyLeuProGlySerProGlyAla	500
QY	797	CCCCAAGCTGATCATGAAGCATAGACAGTAGGCTGGCAGCAGTGT---CCGGGT	741
DB	501	Pro-----GlyThrProGlyProGlnGlyLeuProGly	511
QY	740	---CCCGGAAGAGTCAGAGAGCGGCTCCAGTGGAGTGAAGCAGACCTGGCCACAGA	684
DB	512	SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly-----	529
QY	683	AGTCCAGCAGCCCGCCAGGATGACAGTGCAGCTCCAGGGGCTGGATCCCGGC	624
DB	530	AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla-----ProGly	547
QY	623	ACACAGCCCTGTAGCCAGCGGCGCTGGATGACAAAGAGCTCAGCAGGATGCCCA	564
DB	548	ThrProGly-----ProGlnGly---LeuProGlySerProGlyAlaPro	561
QY	563	AGGACAGTG---CCAGATGAAGGCGGCGGCGGCGCATAGCGTCCACGCGGTGTCAC	507
DB	562	GlyThrProGlyProGlnGlyLeuProGly-----	571
QY	506	TGGTGTAGCTTAGGAGCGGACACAGACAGCCAGCGCCAGCTGGACCAATGCCAGCACCA	447
DB	572	-----SerProGlyAlaProGlyThrProGlyProGln-----GlyLeuPro	585
QY	446	TGGTCACTGAATCTCTCTCTACCCCA-----CTTCAGCAGCAGGAGCGGCA	399
DB	586	GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla	605

D	b		606	-	-----ProGlyThrProGlyProGlnGlyLeuProGlySerProGly	619
Q	y		338	A	GAGCTGGCGTTCCTCCGGT---GCCGCACAGCGGGCTCACCCACAGCCTCTGGACCATAG	282
D	b		620	A	laProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-pr	639
Q	y		281	T	GGGCGCAGGCGGG---TAGGGCTCAGGGGGCCCTCAGGCACCTCCAGAAGCTCCTTCGTCT	225
D	b		639	O	GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe	659
Q	y		224	C	GGCTCTGCTCCAGAAGCTCGGCGCTCTCCTCTCTGCTGCCGCCAACACTCCCTAGGAATCA	165
D	b		659	u	ProGlySerProGlyAlaProGlyThrProGly-----ProGlnGlyLe	674
Q	y		164	G	CCAGGCGCCCATTTCTGCCAGCGCTTGCTGCCGCTTCAGCTTCTCAGCCCATGCTCA-	106
D	b		674	u	ProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP	693
Q	y		105	-----	ACACCTGCTGCTGCTGGGGCACCTCAGTGGGACAGCTGCTCATCACT	60
D	b		693	r	OglyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG	712
Q	y		59	C	AGATCCTGCG---CGAGGCGCGCGCTGTCACCGCGAGCC	22
D	b		712	l	yThrProGlyProGlnGlyLeuProGlySerProGlyAla	725
RESULT 36						
US-08-642-255-101						
; Sequence 101, Application US/08642255						
; Patent No. 5773249						
; GENERAL INFORMATION:						
; APPLICANT: CAPPELLO, Joseph						
; APPLICANT: FERRARI, Franco A.						
; TITLE OF INVENTION: High Molecular Weight Collagen-Like						
; TITLE OF INVENTION: Protein Polymers						
; NUMBER OF SEQUENCES: 135						
; CORRESPONDENCE ADDRESS:						
; ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT						
; STREET: 4 Embarcadero Center, Suite 3400						
; CITY: San Francisco						
; STATE: California						
; COUNTRY: USA						
; ZIP: 94111-4187						
; COMPUTER READABLE FORM:						
; MEDIUM TYPE: Floppy disk						
; COMPUTER: IBM PC compatible						
; OPERATING SYSTEM: PC-DOS/MS-DOS						
; SOFTWARE: PatentIn Release #1.0, Version #1.30						
; CURRENT APPLICATION DATA:						
; APPLICATION NUMBER: US/08/642,255						
; FILING DATE:						
; CLASSIFICATION: 435						
; ATTORNEY/AGENT INFORMATION:						
; NAME: ROWLAND, Bertram I.						
; REGISTRATION NUMBER: 20,015						
; REFERENCE/DOCKET NUMBER: A55556-3/BIR						
; TELECOMMUNICATION INFORMATION:						
; TELEPHONE: (415) 494-8700						
; TELEFAX: (415) 494-8771						
; INFORMATION FOR SEQ ID NO: 101:						
; SEQUENCE CHARACTERISTICS:						
; LENGTH: 837 amino acids						
; TYPE: amino acid						
; STRANDEDNESS: single						
; TOPOLOGY: linear						
; MOLECULE TYPE: protein						
US-08-642-255-101						
Alignment Scores:						
Pred. No.: 2,61e-19 Length: 837						
Score: 356.50 Matches: 263						
Percent Similarity: 35.16% Conservative: 32						

DB:	1	Gaps:	57
US-09-759-143-110 (1-3410) x US-08-477-509B-103 (1-837)			
QY	2347	GGGAAACACAGGTGACTGAGTTTATTAGCTCCCAAAACCTTCTCTAGGTGTCTCTCAA	2288
Db	46	GlySerProGly-----AlaProGlyThrPro-----GlyProGlnGly	58
QY	2287	CTAGGAGGCTAGCTGTTAACCTTGAGCTGGGTGATTCACCTCGACAGTCCCGCATTC	2228
Db	59	LeuProGly-----SerProGlyAlaProGlyThrProGlyProGlnGly	73
QY	2227	AGTGCATGGAGCCCTTCTGCTCCCTGATATAGTCCAGACTGAACCCCTTGGAGGC	2168
Db	74	LeuProGlySerPro---GlyAlaPro-----GlyThrProGlyPro---GlnGly	88
QY	2167	CTCCAGTCAGGACGCTTAGAGACTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2108
Db	89	Leu---ProGlySerPro-----GlyAlaProGlyThrProGly	100
QY	2107	GTGACGCTAGCCACCTCAGCAGCACAGGTGGCAGCAGAGAGAGAGAGAGAGAGAG	2048
Db	101	ProGln-----GlyLeuProGlySerProGlyAlaProGlyThrProGly--P	116
QY	2047	AACAGAAACTGGCGGCAGCCCGGAGCCCTGAGGCTAACAGGAGGAGGAGGAGG	1988
Db	116	roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu	136
QY	1987	CCCAGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1928
Db	136	lySerProGlyAlaProGlyThrPro-----	144
QY	1927	CAAGTCGCTCTTGTCAATATCTACTGTGTAGCAAACTAATGGCAGCAGAGAGG	1868
Db	145	-----GlyProGlnGlyLeuPro-----GlySerProGlyA	155
QY	1867	TGCGGAGAGACACCATATAGGAGTGCAGAGAGTGCAGTGGAGCAATGAGCCAT	1808
Db	155	laProGlyThrProGlyProGlnGlyLeuProGly-----SerPro---	168
QY	1807	CAGGATGGGCGCCACCTGGGAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG	1769
Db	169	-----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly	186
QY	1768	CAGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1721
Db	186	roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr	206
QY	1720	CTCGGTG-----GGCTCAGCCACCCAGCAGCAGCAGCAGCAGCAGCAGCAG	1674
Db	206	roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGly	226
QY	1673	AGGCCCCGAGAGCGCGGTGGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGG	1614
Db	226	lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly	246
QY	1613	CATTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1554
Db	246	roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGly	264
QY	1553	CAGTCTAGCAGCTCAGTGTCCCTCGGTATT-----TGGGAGGAGGAGGAGG	1500
Db	265	Gly-----ProGlnGlyLeuProGlySerProGlyAlaProGlyThr	278
QY	1499	CCCGGTGGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1440
Db	279	ProGlyProGlnGlyLeuPro-----GlySerProGlyAla-ProGlyThrPr	294
QY	1439	CGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1382
Db	294	oGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPr	309
QY	1381	-GGCAGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1323

Db 674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693  
QY 105 -----ACACCTGCTGCTGCTGGGACCTCAGTGGGAGACAGCTCTCATCACT 60  
Db 693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712  
QY 59 CAGATCTGGG---CGAGGCGCGGCTGTCTCACCAGGAGCC 22  
Db 712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725

RESULT 35  
US-08-477-509B-103  
Sequence 103, Application US/08477509B  
Patent No. 5770697  
GENERAL INFORMATION:  
APPLICANT: Ferrari, Franco A  
APPLICANT: Cappello, Joseph  
APPLICANT: Crissman, John W  
APPLICANT: Dorman, Mary A  
TITLE OF INVENTION: No. 5770697el Peptides Comprising Repetitive  
TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,509B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,155  
FILING DATE: 29-DEC-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/053,049  
FILING DATE: 22-APR-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/114,618  
FILING DATE: 29-OCT-1987  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 06/927,258  
FILING DATE: 04-NOV-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Treacartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-55186-7/RFT/MTK  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 837 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-477-509B-103

Alignment Scores:  
Pred. No.: 2,61e-19  
Score: 356.50  
Percent Similarity: 35.16%  
Best Local Similarity: 31.35%  
Query Match: 5.73%

Length: 837  
Matches: 263  
Conservative: 32  
Mismatch: 327  
Indels: 218



QY 2167 CTCAGTCTAGGAGCCCTAGAGACTGGGGAGAGAGAGGAGGAGCCGCCAGCCAGCT 2108  
DB 89 Leu---ProGlySerPro-----GlyAlaProGlyThrProGly 100  
QY 2107 GTGAGCTACGACCTCAGCAGCACAGGTTGGGAGCAGAGAGCCATTTGGCAGC 2048  
DB 101 ProGln-----GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116  
QY 2047 AACAGAAACTGGCGGAGCCCGGAGCCCATCTGGGCTAACAGGAGCGGGAGCTGGGA 1988  
DB 116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136  
QY 1987 CCCAGTACGAGCGCCCTCCACCCCAATGTGCTGGAAGTTTCTACGCTGAGTATTGGC 1928  
DB 136 lySerProGlyAlaProGlyThrPro----- 144  
QY 1927 CAAGTCTGCTTCTCAAACTACTACCTGTGTAGCAAGTAAGTGGGACCCAGCCAGGCC 1868  
DB 145 -----GlyProGlnGlyLeuPro-----GlySerProGlyAla 155  
QY 1867 TGGCGGAGCAGACCATATAGCAGTACAGCTGAGCTGGCAATGGAGCCCATAAA 1808  
DB 155 laProGlyThrProGlyProGlnGlyLeuProGly-----SerPro----- 168  
QY 1807 CAGGATGGGCGCCACCTGGGAGCAGCAGGAGGCAC-----TATC 1769  
DB 169 -----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186  
QY 1768 CAGGATGGGCGGCTCCAGCAGATGCCCGGC-----CCGGAACCCAGCTGCC 1721  
DB 186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206  
QY 1720 CTCGGTG-----GCTCACCACCACCCACACCTACCGAGACATCACAGCA-----G 1674  
DB 206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226  
QY 1673 AGCCCGGAGAGCGGGTGGAGTGGGAGCAGGCGCCACTGTCTCCAGCAGCCAGCTGTC 1614  
DB 226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246  
QY 1613 CATTAGGGAAGGAGCTCCAGCTTAGGCGCTGGCAGGAAGCTGTCTATCAGGCTGCTC 1554  
DB 246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264  
QY 1553 CACTGCTAGCACCCTCCAGTGTCCCTCGGTATT-----TGGCAGGAGAACCTGCTCT 1500  
DB 265 Gly-----ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278  
QY 1499 CCCGTGGTAGAGGAGGCCAGTGTGTAGGCGAGATCTGAGGCGCTGAGAGGTGAACC 1440  
DB 279 ProGlyProGlnGlyLeuPro-----GlySerProGlyAla-ProGlyThrPr 294  
QY 1439 CGGTGAGGCGGCTGAAGCTGTACCCAGGCGCCACTGTGGGACAGGATGTGCGACC-- 1382  
DB 294 oGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPr 309  
QY 1381 -GGCAGCCACAGGGAAGCTGCCACTGTGCGCAAAATAGACTGCTCGAGTCCGGAATCGCT 1323  
DB 309 oGly-ProGlnGly--LeuProGlySerPro-----Gly 319  
QY 1322 GCACAGCGGTTCATGACAGAGAGAGACAGGAGATGGCGCACTGAGGAACA--- 1266  
DB 320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339  
QY 1265 -----GCCCGAGCTGCCATCC-----GACGCGCTT 1239  
DB 340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359  
QY 1238 CATCATAGTGTCTCCGGGCTCGGTGCGCGGCTCAGCTCTGGGACGCCCTGTGTACAGCC 1179  
DB 360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

QY 1178 CCTCCGCCAGCAATCCGTGTAAACACAGCGTGAAGGTCAATGAGTCCCATCCAGC----- 1125  
DB 379 GlyAlaProGlyThrProGlyProGlnGly-----LeuProGlySerProGly 394  
QY 1124 -----TGCACAGCTCAGCCAGCAGAGAGCCGCGAGGTGGGGAGTCCGGCAGCACA 1071  
DB 395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414  
QY 1070 GCT-----GGTGCAGCCGGGAAAGAGGCGCCAGTTCCTCGGAAAGCCAGCGGG--- 1020  
DB 415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434  
QY 1019 CCCGCGATGACAGCAGTGGGCGGACAGAGAGGAGGGGGCGCAGCCCTTCGCTGCTCGG 960  
DB 435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453  
QY 959 TGGGCGCCAGCGCTCCCT---CCTCAGCCACAGCAGTGTGGCTGTACGAGGTGAGGA 903  
DB 454 GlyAlaProGlyThrProGlyProGln-----Gly 463  
QY 902 AGATGAGGTGAGCAGGCGCCAAAGAGGACTCCT-----CCTGGGTGCCCA 858  
DB 464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483  
QY 857 GGTAGGGGGCGGCGCAGCTGTGTCCAGTCAATGGCAGGAGGAGTACCCAGCAGC 798  
DB 484 Gly---AlaProGly-----ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500  
QY 797 CCCCAAGACTGATCATCAAGCATAGACAGTAGGCTGCGGACAGTGT---CCGGGT 741  
DB 501 Pro-----GlyThrProGlyProGlnGlyLeuProGly 511  
QY 740 ---CCCGAAGAGTCAAGAGCAGGAGGCTCCAGTGGAGTGAAGCAGCAGCTGCCACAGA 684  
DB 512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly----- 529  
QY 683 AGTCCAGCAGCCAGCCAGGATGAGCAGTGCAGCTCCAGGCGCTGGGATCGGGC 624  
DB 530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla-----ProGly 547  
QY 623 ACAGCAGCCCTGTAGCAGCGCGCTTGGGATGAGAAAGAGCCTCAGCAGGATGCCCA 564  
DB 548 ThrProGly-----ProGlnGly---LeuProGlySerProGlyAlaPro 561  
QY 563 AGGACAGTG---CCCAGATGAAGGCGCGCGGCGCATAGCTCCACGCGCTGCTAC 507  
DB 562 GlyThrProGlyProGlnGlyLeuProGly----- 571  
QY 506 TGGCTGAGCTTAGGCGGAGCAGACAGCAGGCGCCAGCAGTGGACCAATGCCAGCACA 447  
DB 572 -----SerProGlyAlaProGlyThrProGlyProGln-----GlyLeuPro 585  
QY 446 TGTCTATGAATCTCTCTTACCCCA-----CTTCCAGCAGCAGAGCGGCA 399  
DB 586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605  
QY 398 CATAGTGTATGCTCGGCGCAACACACCTCCAGGCGCAAGGTAGCAGTTGACAGCA 339  
DB 606 -----ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619  
QY 338 AGAGCTGGGCTTTCGGT---GCCGAGCAGGCGGCTCACCCAGCAGCTCTGACCATAG 282  
DB 620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639  
QY 281 TGGCGCAGCGGG---TAGGGCTCAGGGGCGGCTTCCAGGCACTCCAGAACTCTCTTCGCT 225  
DB 639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659  
QY 224 CGGCTCTGCTCCAGAGCTGCGGCGCTCTCTCTGCTGCCGCCAACTGCCTAGGAATCA 165  
DB 659 uProGlySerProGlyAlaProGlyThrProGly-----ProGlnGlyLe 674  
QY 164 GCCAGGCGCCCATTTCTGCCAGCCCTTGGTCCCGGCTCCAGCTTCTCAGCCCATGCTCA- 106

QY 95 CTGTGGGGACCTCAGTGGGGACAGCTCTCATCTCAGATCCTGCGC---CGAGGGCGCGC 39  
Db 679 roGlnGly---LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuP 698  
QY 38 GGCTGTCCACCGGAGCC 22  
Db 698 roGlySerProGlyAla 703  
RESULT 34  
US-08-175-155-68  
; Sequence 68, Application US/08175155  
; Patent No. 5641648  
; GENERAL INFORMATION:  
; APPLICANT: Ferrari, Franco A.  
; APPLICANT: Cappello, Joseph  
; APPLICANT: Crissman, John W.  
; APPLICANT: Dorman, Mary A.  
; TITLE OF INVENTION: Methods for Preparing Synthetic  
; TITLE OF INVENTION: Repetitive DNA  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/175,155  
; FILING DATE: 29-DEC-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I.  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: A-55186-5/BIR  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 837 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-175-155-68  
Alignment Scores:  
Pred. No.: 2,61e-19 Length: 837  
Score: 356.50 Matches: 263  
Percent Similarity: 35.16% Conservative: 32  
Best Local Similarity: 31.35% Mismatches: 327  
Query Match: 5.73% Indels: 218  
DB: 1 Gaps: 57  
US-09-759-143-110 (1-3410) x US-08-175-155-68 (1-837)  
QY 2347 GGGAAACACAGGTGACTGAGTTATTTCAGCTCCCAAAACCTTCTTAGTGTGTCTCAA 2288  
Db 46 GlySerProGly-----AlaProGlyThrPro-----GlyProGlnGly 58  
QY 2287 CTAGGAGGTAGCTGTTACCCGAGCCTGGGTATTCACCTGCAGAGTCCCGCGCATTC 2228  
Db 59 LeuProGly-----SerProGlyAlaProGlyThrProGlnGly 73  
QY 2227 AGTGCATGGAGCCCTTCTGCCTCCCTCTGATATAGTCCAGACTGAAACCCCTTGGAGGC 2168  
Db 74 LeuProGlySerPro---GlyAlaPro-----GlyThrProGlyPro---GlnGly 88

QY 1106 AGACCGGGCGCAGGTGCGGGCATGCGCAGACAGCT-----GGTGCAGCCGGGAA 1053  
Db 381 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 400  
QY 1052 GCAGGGCGCCAGGTTCGGAAAGCCAGCGG---CCCGCATGGACAGCAGTGGGCGC 996  
Db 401 ProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGly 420  
QY 995 ACAAGAGGGGGCCAGACCCCTTCTGCTGGCTCGGTGGGCGCCAGCGCTGCCT---CCT 939  
Db 421 Thr---ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 439  
QY 938 CAGCCACCGACAGTGGCTGCTACGCGAGTGAAGATCAGGCGTGAAGCCAAAGA 879  
Db 440 Gln-----GlyLeuProGlySerProGlyAlaPro 449  
QY 878 GGCATCTCT-----CCTGGTCCCGAGGTAGGGGCGCCAGCGCTGGTGT 834  
Db 450 GlyThrProGlyProGlnGlyLeuProGlySerProGly---AlaProGly-----Thr 466  
QY 833 CCCAGTCAATGGCAGGAGGTAGCCAGCAGCCCAAGACTGATCATGAAGCAT 774  
Db 467 ProGlyProGlnGlyLeuProGlySerProGlyAlaPro----- 479  
QY 773 AGACAGAGTAGGCTGGCGACAGTGGT---CCGGGT---CCCGAAGAGGTGACAGACA 720  
Db 480 -----GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 497  
QY 719 GGGCTTCAGTGAAGTGAAGCAGCAGTCCGCCAGAGTCCAGCAGCCCGCCAGGA 660  
Db 498 GlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrProGlyProGln 515  
QY 659 TGAGCAGTGCACCTCAGGGGCTGGATCGCGGCACAGCAGCCCTGTAGCAGCCGG 600  
Db 516 GlyLeuProGlySerProGlyAla-----ProGlyThrProGly----- 528  
QY 599 CCCTTGGATGAGAAAGAGCTCAGCAGGATGCCCAAGGACAGTG---CCCAGATGAAGG 543  
Db 529 ProGlnGly---LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 547  
QY 542 CCGCGGGCGGCTAGCTGCGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483  
Db 548 ProGly-----SerProGlyAlaProGly 555  
QY 482 AGACGAGCGCCAGCAGTGGACCAATGCCAGCAGCAGTGTGCTGCTGCTGCTGCTGCT 423  
Db 556 ThrProGlyProGln-----GlyLeuProGlySerProGlyAlaProGlyThr 571  
QY 422 CCA-----CTTCAGCAGCAGAGCGCGGCACATAGTGTGCTGCTGCTGCTGCTGCTGCT 375  
Db 572 ProGlyProGlnGlyLeuProGlySerProGlyAla-----ProGly 585  
QY 374 ACACCTCCAGCGCCAAAGTTAGCAGGTGACCAAGAGAGCTGGGCTTTCCGGT---GCC 318  
Db 586 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGln 605  
QY 317 GCAGCAGCGGGCTCAGCAGCAGCTCTGAGCAGCAGTGTGCGCGCGCGGG---TAGGGCTC 261  
Db 606 GlyLeuProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGly 625  
QY 260 AGGGGGCGGTTCAGGCACTCCAGAACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201  
Db 625 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGly 645  
QY 200 CTCCTCTCTTCTGCTGCGCCAACTGCTAGGAATCAGCGCGCGCGCGCTTCTGCGCAGCC 141  
Db 645 yThrProGly-----ProGlnGlyLeuProGlySerProGlyAla---Pr 659  
QY 140 CTTTGGTGGCTCCAGCTTCTCAGCCCATGCTCA-----ACACCTGCTG 96  
Db 659 oGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProGlyThrProGly 679





```

Qy 39 CGGCTGTACCCGGA 25
    |||
Db 693 ArgGlyAspProGly 697

RESULT 32
US-08-642-255-132
; Sequence 132, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 829 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-255-132

Alignment Scores:
Pred. No.: 2,38e-19 Length: 829
Score: 357.00 Matches: 256
Percent Similarity: 34.78% Conservative: 33
Best Local Similarity: 30.61% Mismatches: 324
Query Match: 5.73% Indels: 219
DB: 1 Gaps: 53

US-09-759-143-110 (1-3410) x US-08-642-255-132 (1-829)
QY 2320 GCTCCAAAACCGCTTCTAGGTGTCTCT---CAACTAGAGAGCTAGC
Db 32 AlaProGlyThrProGlyGluGlyGlnGlnHisLeuGlyGlyAla
QY 2263 AGCCTGGGTAAATCACTGCAGAGTCCCGCATTCACAGTCGACGAGC
Db 52 AspValGlySerProGlyAlaProGlyThrPro-----GlyAlaProGly
QY 2203 CPTGTATAAGTCCAGACTGAAACCCCTTGGAAGCGCTCCAGTCAGGCG
Db 68 Pro-----GlyAlaProGly
QY 2143 TGGGGAGAGAGGAGAGGAGCGCCCGACCCCGACCTGTGCAGCTACGCG
Db 72 -----GlyAlaProGlyThrProGlyProGln-----

```

APPLICANT: CAPPELLO, Joseph  
 APPLICANT: FERRARI, Franco A.  
 TITLE OF INVENTION: High Molecular Weight Collagen-Like  
 TITLE OF INVENTION: Protein Polymers  
 NUMBER OF SEQUENCES: 135  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
 STREET: 4 Embarcadero Center, Suite 3400  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-4187  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/642,255  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ROWLAND, Bertram I.  
 REGISTRATION NUMBER: 20,015  
 REFERENCE/DOCKET NUMBER: A55556-3/BIR  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 494-8700  
 TELEFAX: (415) 494-8771  
 TELEX: 910 272299 FHT UR  
 INFORMATION FOR SEQ ID NO: 62:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1064 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-642-255-62

Alignment Scores:  
 Pred. No.: 1,53e-19 Length: 1064  
 Score: 360.00 Matches: 265  
 Percent Similarity: 33.26% Conservative: 24  
 Best Local Similarity: 30.49% Mismatches: 296  
 Query Match: 5.78% Indels: 286  
 DB: 1 Gaps: 60

US-09-759-143-110 (1-3410) x US-08-642-255-62 (1-1064)  
 QY 2314 AAAACCCCTCTCTAGGTGTCTCACTAGGAGGCTAGCTGTTAAACCTGAGCCTGGGT 2255  
 Db 12 GluAsnPro-----GlyValThrGln-----LeuAsnArgLeuAlaAla 24  
 QY 2254 AATCCACCTGCAGAGTCCCGCGATTCCAGTGCATGCAGGCCCTCTGGGCTCCCTGTATAA 2195  
 Db 25 HisProPheAlaSerAspProMetGlyAlaHisGlyProAlaGly----- 40  
 QY 2194 GTCCAGACTGAACCCCTTGAAGGCTCCAG-----TCAGGAGCCCT 2150  
 Db 41 -----ProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaPro 54  
 QY 2149 AGAGACTGGGAGAG-----AGGAGAGGAGCCGCCAGCCCGCCAGCTGTCAGCTACGC 2096  
 Db 55 GlyProAlaGlyProProGlySerArgGlyAspProGlyPro----- 69  
 QY 2095 ACCTCAGCAGCAGCGGTGGCAGCAGAGAGCCATTACTTTGGCAGCAACAGAACTGG 2036  
 Db 70 -----GlyAlaProGlyProAlaGlyProPro-----GlySerArgGlyAspPro 84  
 QY 2035 CGGCCAGCCCGCAGCCCGCTACAGAGAGCGGGAGCTGGGA----- 1988  
 Db 85 GlyProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyPro 104  
 QY 1987 -----CCNAGTCAGCAGCCCTCCACCCCATGTGCTGGAGTGTCTACGCTGA 1937

Db 105 ProGlyAlaProGlyProAlaGlyProPro-----GlySer----- 116  
 QY 1936 GTATTTGGCCNAGTCGCTCTGTCAATACTACTCTGTGTAGCAAAAGTAAATGGCAGCAG 1877  
 Db 117 -----ArgGlyAspPro 120  
 QY 1876 AC-----CCAGCGCTCGCGCAGACACCATATAGGCAGTGCACAGATGGCGTGA 1830  
 Db 121 GlyProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyPro 140  
 QY 1829 GCTGGACAATGGAGCCCAATACAGGGATGGGCCACCTGGCAGCAGCAGCA-----AGGCAC 1773  
 Db 141 ProGlyAlaProGlyPro-----AlaGlyProProGlySerArgGlyAspProGly 157  
 QY 1772 TATCCAGGATGGCGAGGTCCAGGCAGATGCCCGGC-----CCGGAACCAACC 1725  
 Db 158 ProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyPro 177  
 QY 1724 TGG---CCTCGGTGGGCTCACCCACCACC-ACAGGTACGGAGACATCACAGGAGAGGCC 1669  
 Db 178 GlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyPro 197  
 QY 1668 CCGCAGAGCGGGGTGGAGTGGGAGCAGGCCACTGCTCCAGCAGCCACGTCCTCATTA 1609  
 Db 198 ProGlyProAlaGlyProProGlySerArgGlyAspProGlyProProGlyAlaProGly 217  
 QY 1608 GGGAAAGGAGCTCCAGGCTTA-----GGGCGTGGCAGGAAGCTGGTGCATCAGGCTGTC 1555  
 Db 218 ProAlaGlyProProGlySerArgGlyAspProGly----- 229  
 QY 1554 TCACTGTAGCACCTCAGCTGTCCTCGGTATTTGGGAGGAGAACACCTGCTCTCCCGG 1495  
 Db 230 -----ProProGlyAlaPro-GlyPro---AlaGlyProProGlySerArg 244  
 QY 1494 TGGTAGAGGGAGGCGCTGTCT-----AGGGCAGGATCTCCAGGCTGAGAGGTGAAC 1441  
 Db 244 YAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySerArg-GlyAsp 264  
 QY 1440 CCGGTGAGGGCGGTGAAGCTGTCCACCCAGC-----CACACTG 1402  
 Db 264 roGlyProProGly-----AlaHisGlyProAlaGlyProLysGlyAlaHis----- 279  
 QY 1401 TGGGACAGGATGTGGCAGCGGAGCCACACAGG-----AAAGTGCACACTG 1354  
 Db 280 --GlyProAlaGlyProLysGlyAlaHis-GlyProAlaGlyProLysGlyAla-HisG 298  
 QY 1353 GCCAAATAGACTGCTCGAGTCCGAATCGCTGCACACCGCGT---CCATGACCAGAGAG 1297  
 Db 298 yPro-----AlaGlyProLysGlyAlaProGlyProAlaG 310  
 QY 1296 AAGACCGAGAGATGGCGCACTGCAGGAACAGCCCGAGGCTGCCATCCGAAACGCTTCA 1237  
 Db 310 yProProGlySerArgGlyAspProGly---ProProGlyAlaPro----- 324  
 QY 1236 TCATAGTGTCTCCGGGCTCGTCCCGGCTCAGCTGCGCAGCAGCCCTGGTACA----- 1182  
 Db 325 -----GlyProAlaGlyProProGlySerArgGlyAspProGlyPro 340  
 QY 1181 -GCCCTCGCCACGAAATCGTGTAAACACGCGTGAAGGTGATGAGTCCCATCAGCTG 1123  
 Db 340 yAlaProGlyProAlaGlyPro-----ProGlySer-- 350  
 QY 1122 CACAGCTCAGCCACGAAGAGCGCGCA-----GGTGGCGGCATCGCGCAGCAGCAGC 1069  
 Db 351 -ArgGlyAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySerArg 370  
 QY 1068 T-----GGTGCAGCGGGAGCAGGCGCCCGCTCC----- 1035  
 Db 370 yAspProGlyProProGlyAlaProGlyProProGlySerArgGlyAspPro 390  
 QY 1034 -GGAAAGCAAGCGGGCGCGCATGGACAGTGGGCGCAGCAAG-----AGGGGCGC 982





Db	657	--GlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProG	676
Qy	1969	-----GAGGGCGCTCCCTCACTGGGTCCCGACAGCTCCCGCTCCTGTAGCCCC	2014
Db	676	lyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrProGlyPro	695
Qy	2015	ATGGGGCTCGCGGCTGGCGG	2035
Db	696	GlnGlyLeuProGlySerPro	702
RESULT 30			
US-08-397-633A-31			
Sequence 31, Application US/08397633A			
Patent No. 5773577			
GENERAL INFORMATION:			
APPLICANT: Capello, Joseph			
TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE			
TITLE OF INVENTION: OF ENZYMATIC CROSS-LINKING			
NUMBER OF SEQUENCES: 105			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT			
STREET: 4 Embarcadero Center, Suite 3400			
CITY: San Francisco			
STATE: California			
COUNTRY: USA			
ZIP: 94111-4187			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/397,633A			
FILING DATE:			
CLASSIFICATION: 530			
ATTORNEY/AGENT INFORMATION:			
NAME: Rowland, Bertram I			
REGISTRATION NUMBER: 20,015			
REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (415) 781-1989			
TELEFAX: (415) 398-3249			
TELEX: 910 277299			
INFORMATION FOR SEQ ID NO: 31:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 762 amino acids			
TYPE: amino acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: peptide			
US-08-397-633A-31			
Alignment Scores:			
Pred. No.:	6.7e-20	Length:	762
Score:	364.00	Matches:	223
Percent Similarity:	33.95%	Conservative:	32
Best Local Similarity:	29.69%	Mismatches:	300
Query Match:	5.67%	Indels:	197
DB:	1	Gaps:	43
US-09-759-143-110 (1-3410) x US-08-397-633A-31 (1-762)			
Qy	84	GGTCCCCACAGCAGCAGGTGTG-----AGCATGGCTGAGAACTGGACGGCACCA	137
Db	52	GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro	71
Qy	138	AAGGGC-----TGGCAGAAATGGCGCCTGGC-----TGATTC	170
Db	72	GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlnGlnHis	91
Qy	171	CTAGGCAGTTGGCGGCACGAAGGAGGAGCGCGCAGCTTCTGGACGACGCGGACCA	230
Db	92	LeuGlyGlyAlaArg---GlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro	110



Db 214 -----SerAsnGlyAlaPro-----GlyGlnAr 221  
QY 2129 AGGAGCGCCAGCCCGCCAGCTGTGCAGCTACGCACACAGGCTGGCAGCAG 2070  
Db 221 gGlyGluProGlyProGlnGly-----HisAlaGlyAlaGlnGlyProProGly 237  
QY 2069 AGAGCCACATTTTGGCAGCAACAGAAACTGGCGCCAGCCCGCAGCCCGCCATGGGC 2010  
Db 237 yProProGlyTleAsnGlySer-----ProGlyGly-- 247  
QY 2009 TAACAGGAGCGGGAGCTGGGACCCAGTGGAGCGCCCTCCACCACCAATGTCTGGAAG 1950  
Db 248 -----LysGlyGluMetGlyProAla--GlyTleProGlyAlaPro----- 260  
QY 1949 TTTTCTAGCTGAGTATTTGGCCAACTGGCTCTGTCAATACTACCTGTGTAGCAAAAGT 1890  
Db 261 -----GlyLeuMetGlyAlaArgGly-----ProProGlyProAla 272  
QY 1889 AAATGGCGACAGACCCAGCTGGCGCAGACACCATATAGGCAGTACAGACTGGCTGA 1830  
Db 273 GlyAlaAsnGlyAlaProGlyLeuArg-----GlyGly 283  
QY 1829 GCTGGACAATGGACCCCAATAACAGGATGGGCCACCTGGGACAGCAGCAAGGCAC-- 1773  
Db 284 AlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyProArgGlyGluArgGly 303  
QY 1772 -----TATCCAGGATGGCGAGGT-----CCA 1752  
Db 304 GluAlaGlyTleProGlyValProGlyAlaLysGlyGluAspGlyLysAspGlySerPro 323  
QY 1751 GGCAGATGCCCGCCGCGGAACACCTGGCTCGGTGGGCTCACCCACCACACAGCTA 1692  
Db 324 Gly-AspProGlyAlaAsnGlyLeuProGlyAlaAlaGly-----GluAr 338  
QY 1691 CGGAGACATCACAGCAGAGCCCGCA-----CAGCGCGGTG 1653  
Db 338 gGlyAlaLeuGlySerArgGlyProAlaGlyProAsnGlyTleProGlyGluLys-Glyp 358  
QY 1652 GAGTGGGAGAGCCCACTGCTCCAGCACCCAGCTGCTCCATTAGG--AAGGGAGCTC 1596  
Db 358 roAlaGlyGluArgGly---AlaProGlyProAlaGlyProArgGlyAlaAlaGlyGluP 377  
QY 1595 CAGGCTTAGG--CCTGGCAGGAGCTGGTGCATCAGGCTGCTCCTACTGCTAGCAC 1542  
Db 377 roGlyArgAspGlyValProGly-GlyProGlyMetArgGlyMetPro----- 392  
QY 1541 CTCCAGTCTCCCTCGGTATTGGGCAGCAACACCTGCTCTCCCGTGTGTAGAGGAGG 1482  
Db 393 -----GlySerProGlyGlyProGlySerAspGlyLys 403  
QY 1481 CCAGTGTGTAGGCGAGATCTGCAGGCTGAGAAGTGAACCCGGTGAAGGCGGCTGAAG 1422  
Db 404 Pro---GlyProProGlySer-GlnGlyGluSerGlyArgProGlyProProGlyProSe 422  
QY 1421 CTGTCACACCGCCACACTGTGGCAGAGGATGTGGCACCAGCGGCACAGCAAGAACTG 1362  
Db 422 r--GlyProArgGlyGlnProGlyVal---MetGlyPheProGlyProLysGlyAsn-- 439  
QY 1361 CCACACTGGCCAAATAGACTGCTCAGTGGCGGAATCGCTGCACCAGCGGTTCATGACCA 1302  
Db 440 AspGlyAlaProGlyLysAsnGlyGlu-----ArgGlyGlyProGlyProGlyPro 457  
QY 1301 CAGAGAAGACCGAGATGGCGCACTGCAGGA-----ACAGCC 1263  
Db 458 GlnGlyProProGlyLysAsnGlyGluTyGlyProGlnGlyProProGlyProThrGly 477  
QY 1262 CCAGCTGCCCATCCGAACGCTTCATCATAGTCTCCGGGCTCGGTCCCGGCTCAG 1203  
Db 478 ProGlyGlyAspLys-----GlyAspThrGlyProArgGlyProGlnGly 492  
QY 1202 CTCTGGGCACCCCTGGTACA----- 1182  
Db 493 LeuGlnGlyLeuProGlyThrGlyGlyProProGlyGluAsnGluLysProGlyGluPro 512

1181 -----GCC 1179  
513 GlyProLysGlyGluAlaGlyAlaProGlyAlaProGlyLysGlyAspAlaGlyAla 532  
1178 CTTCCGCCACAGAAATCCGTGTAAACAGCGTGAAGTCAATGATCCATCCAGCTGCACA 1119  
533 ProGlyGluArgGlyPro-----ProGlyLeuAlaGly 543  
1118 GCTACCCACAGAGCGCGCGCAGGTCGGGGCATCGGCAGCAGACAGCTGTGCAGCC 1059  
544 AlaProGlyLeuArgGlyAlaGlyProProGlyProGlyGlyLysGlyAlaAla 563  
1058 GGGAGAACAGCGCGCCAGGTTCC----- 1035  
564 GlyProProGlyProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGlu 583  
1034 -----GGAAAGCCAGCGCGCCGCC 1014  
584 ArgGlyGlyLeuGlySerProGlyProLysGlyAspLysGlyGluProGlyProGly 603  
1013 ATGGACAGCAGTGGGCGACA-----AGGAGGGCGCGCAGACGCCCTTCTGCTGGCTCGG 960  
604 AlaAspGlyValProGlyLysAspGlyProArgGlyProThr----- 617  
959 TGGGCGCCAGCGCTGCTCTCAGCCACAGCAGTGTGCTGCTAGCAGGTGAGGAAGA 900  
618 --GlyProIleGlyProProGlyProAla-----GlyGln 628  
899 TGAGGTGAGCAGGCCAAAGAGCAGCTCTCTCTGGTGGCCAGGT---AGGGGGCCAGGG 843  
629 ProGlyAspLysGlyGlyGlyAla---ProGlyLeuProGlyTleAlaGlyPro--- 646  
842 CACTGTGTCCCACTCAATGGCAGGAGGTAGCCAGCAGCGCCCAAGACTGATCA 783  
647 -----ArgGlySerProGlyGlu----- 552  
782 TGAAGCATAGACAGAGTAGGCTGGCAGACAGTGT---CCGGT---CCCGAAGAGGT 729  
653 --ArgGlyGluThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsn 671  
728 CAGAGAGCAGGCGCTCCAGTGGAGTGAACACACACCTGGCCACAGAGTCCAGCAGCCCA 669  
672 GlyGluProGlyLysGlyGluArgGlyAlaProGlyGlyLysGlyGlyGlyPro 691  
668 CGCCAGGATAGCAGTGCAGCTCCAGG-----GCCTGGGATCCGGCCAGCAGA 618  
692 ---ProGlyValAlaValProProGlyGlySerGlyProAlaGlyProGlyProGln 710  
617 GCCTGCTAGCCAGCGCGCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCAAGGACA 558  
711 GlyValLysGly-----GluArgGlySerProGlyGlyPro----- 722  
557 GTCCCCAGATGAAGGCGCGCGCCCATAGCTCCACCGCATAGTCTGCTGCTGCTGAGC 498  
723 -----GlyAlaAlaGlyPhe----- 727  
497 CTAGAGCGGACACAGACAGCGCCAGCAGCTGGACCAATGCCAGCAGCAGTGTGATCA 438  
728 ProGlyAlaArgGlyLeuProGlyPro-----ProGlySerAsn 740  
437 ACTTCTCTTACCCCACTTCCAGCAGCAGAGCGCGGCACATAGTGTGCTGCGGCCA 378  
741 GlyAsnProGlyProProGlyProSerGlySer----- 751  
377 AACACACTCCAGGCCAAAGGTTAGCAGGTTGACACAGCAAGAGCTGGGCTTCCCGTGC 318  
752 -----ProGlyLys-----AspGlyProProGlyProAlaGlyAsnThrGlyAla 766  
317 GCAGCAGGCGGCTCACCCACAGCAGCTCTGGACCATAGTGGG-----CCAGGCGG 270  
767 Pro-----GlySerProGlyValSerGlyPro-LysGlyAspAlaGlyGlnProGlyGln 784



QY	2249	ACCTGCAGAGTCCCGCATTCACAGTGCATGAGGCCCTTCTGCGCTCCCTGTATAGTCCA	2190	QY	1262	CCAGGCTGCCCATCCGAACCCCTTCATCATAGTGTCTCCGGGGCTCGGTGCCGGGTCTAG	1203
Db	200	oGlyAlaLysGlyGluValGlyProAlaGlySerProGly	213	Db	478	ProGlyGlyAspLys	492
QY	2189	GACTGAAACCCCTTTGGAAGGCTCCAGTCCAGGAGCCCTAGAGACTGGGAGAGAGGAG	2130	QY	1202	CTCTGGGCAGCCCTGGTACA	1182
Db	214	-----SerAsnGlyAlaPro-----GlyGlnAr	221	Db	493	LeuGlnGlyLeuProGlyThrGlyGlyProGlyGluAsnGluLysProGlyGluPro	512
QY	2129	AGGAGAGCCCGCCAGCCAGCTGTGCACCTAGCCACCTCAGCAGCAGAGGTGCGCAGCAG	2070	QY	1181	-----GCC	1179
Db	221	gGlyGluProGlyProGlnGly-----HisAlaGlyAlaGlnGlyProProGly	237	Db	513	GlyProLysGlyGluAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAla	532
QY	2069	AGAGCCACATATTCTTGGCAGCAACAGAACTGGCGGCAGCCCGCAGCCCATTTGGGC	2010	QY	1178	CCTCGCCCAACAAATCCGTGTAAACACCGCTGAAGTCAATGAGTCCATCCAGCTGCACA	1119
Db	248	-----LysGlyGluMetGlyProAla--GlyIleProGlyAlaPro-----	260	Db	533	ProGlyGluArgGlyPro-----ProGlyLeuAlaGly	543
QY	1949	TTTTCTACGCTAGTATTGGCCAAAGTCGCTCTTGTCAATACTACCTGTGTAGCAAGT	1890	QY	1118	GCTCAGCAGCAGAGCCGGCGAGGTGCGGGCATGCGGCAGCAGCAGTGTGTGAGCC	1059
Db	261	-----GlyLeuMetGlyAlaArgGly-----ProProGlyProAla	272	Db	544	AlaProGlyLeuArgGlyGlyAlaGlyProProGlyProGlyGlyLysGlyAlaAla	563
QY	1889	AAATGGCGACAGACCCAGCCCTGCGGCAGACACCATATAGCAGCTGACACTGGCTGA	1830	QY	1058	GGGGAAGCAGGCGCCCGAGGTTC	1035
Db	273	GlyAlaAsnGlyAlaProGlyLeuArg-----GlyGly	283	Db	564	GlyProGlyProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGlu	583
QY	1829	GCTGGACATGAGGCCCAATACAGGGATGGGGCCACCTGGGCAGCAGGAGGAC	1773	QY	1034	-----GGAAGCCAGCGGCCCGC	1014
Db	284	AlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyProArgGlyGluArgGly	303	Db	584	ArgGlyGlyLeuGlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGly	603
QY	1772	-----TATCAGGATGGCGAGGT-----CCA	1752	QY	1013	ATGCACAGCAGTGGGGGACA-----AGGAGGGCGGCAGACAGCCCTCTCTGCTGGCTCGG	960
Db	304	GluAlaGlyIleProGlyValProGlyAlaLysGlyGluAspGlyLysAspLysSerPro	323	Db	604	AlaAspGlyValProGlyLysAspGlyProArgGlyProThr-----	617
QY	1751	GGCAGATGCCCGGCCCGGAAACCCCTGCGGTGGCTCAGCCAGCTGACACACGTA	1692	QY	959	TGGGGCCCGAGCGCTGCTCTCAGCCACAGCAGTGTGGCTGTACCGAGTGAGGAAGA	900
Db	324	Gly-AspProGlyAlaAsnGlyLeuProGlyAlaAlaGly-----GluAr	338	Db	618	--GlyProIleGlyProGlyProAla-----GlyGln	628
QY	1691	CGGAGATCACAGGAGGCGCCGCA-----GAGCGCGGTG	1653	QY	899	TGAGGTGAGCAGCGCCAAAGAGCAGCTCTCTCTGGTCCAGCT-----AGGGGCCAGGG	843
Db	338	gGlyAlaLeuLysArgGlyProAlaGlyProAlaGlyProAlaGlyLysGlyLysGly	358	Db	629	ProGlyAspLysGlyGlyGlyGlyAla-----ProGlyLeuProGlyIleAlaGlyPro	646
QY	1652	GAGGTGGAGCAGGCGCTCCCTCCAGCAGCCAGCTGTCCATTAGGG-----AAGGAGCTC	1596	QY	842	CAGTGTCTCCAGTCAATGCGCAGGAGGTAGCCAGCAGCAGCCAGCCAGCTGTATCA	783
Db	358	roAlaGlyGluArgGly-----AlaProGlyProAlaGlyProArgGlyAlaAlaGlyGlu	377	Db	647	-----ArgGlySerProGlyGlu-----	652
QY	1595	CAGGCTTAGGG-----CCTGGCAGGAAGCTGTATCAGGCTGTCTCAGCTGTAGCAC	1542	QY	782	TGAAGCATACAGAGTAGGCTTGGCCAGCAGTGGT-----CCGGCT-----CCCGGAGAGGT	729
Db	377	roGlyArgAspGlyValProGly-GlyProGlyMetArgGlyMetPro-----	392	Db	653	---ArgGlyGluThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsn	671
QY	1541	CTCAGTGTCCCTCGGTATTGGGCGAGGAACACCTGCTTCTCCCGGTGGTAGAGGAGG	1482	QY	728	CAGAGCAGGCGCTCCAGTGGAGTGAAGCAGCAGCTGGCCACAGAGTCCAGCAGCCCA	669
Db	393	-----GlySerProGlyGlyProGlySerAspGlyLys	403	Db	672	GlyGluProGlyGlyLysGlyGluArgGlyAlaProGlyGluLysGlyGlyGlyPro	691
QY	1481	CCAGTGTAGGCGAGGCTGAGGCTGAGAGGTGAACCGGTGAGCGCGCTGAAG	1422	QY	668	CGCCAGATGAGCAGTCCAGCTCCAGGG-----GCCTGGATCCCGGCACAGCA	618
Db	404	Pro--GlyProProGlySer-GlnGlyGluSerGlyArgProGlyProGlyProse	422	Db	692	--ProGlyValAlaValProProGlyGlySerGlyProAlaGlyProProGlyProGln	710
QY	1421	CTGTACAGGCGCACACTGTGGGACAGGATGTGGCCCGCAGCCAGCAGGGAAGCTG	1362	QY	617	GCCCTGTAGCAGCGCCCTTGGATGACAAAGAGCTCAGCAGGATGCCCAAGGACA	558
Db	422	r--GlyProArgGlyGlnProGlyVal---MetGlyPheProGlyProLysGlyAsn	439	Db	711	GlyValLysGly-----GluArgGlySerProGlyGlyPro-----	722
QY	1361	CCACTGGCCAAATAGCTGTCTGAGTGCAGTCCGAATCGCTGCACCGCGGTCCATGACCA	1302	QY	557	GTGCCACAGTGAAGGGCGCGCGGCATAGCGTCCACAGCTGTGTCTGTGCTGAGC	498
Db	440	AspGlyAlaProGlyLysAsnGlyGlu-----ArgGlyGlyProGlyGlyProGlyPro	457	Db	723	-----GlyAlaAlaGlyPhe-----	727
QY	1301	GAGAGAAGACGAGGATGGCGCACTGCAGGA-----ACAGCC	1263	QY	437	CTAGGAGCGGGACACAGACAGCCAGCCAGCTGGACCAATCCCGACACCATGTGTATGA	438
Db	458	GlnGlyProProGlyLysAsnGlyGluTyrgGlyProGlnGlyProProGlyProThrGly	477	Db	728	ProGlyAlaArgGlyLeuProGlyPro-----ProGlySerAsn	740
				QY	437	ACTTCTCTCTACCCCGCTTCCAGCAGCAGGCGGCACATAGTGTATGCTCGGGCA	378
				Db	741	GlyAsnProGlyProProGlyProSerGlySer-----	751
				QY	377	AAACACCTCCAGGCGCAAAAGTTAGCAGTTGACCGCAAGAGCTGGGCTTTCCGGTGCC	318





QY 2363 TAAGGGCTTAGA-----GATGGAAACACAGGTGACTGAGTTATTTCAGCTCCCAAAA 2310  
Db 176 -AspGlyAlaArgGlySerAspGlyGlnProGly-----ProProGly 189  
QY 2309 CCCTTCTCTAGTGTCTCAACTAGGAGGCTAGCTGTTAAACCTGAGCCTGGGTATCC 2250  
Db 189 yProPro---GlyThrAlaGlyPhePro-----GlySerPr 200  
QY 2249 ACCTGAGGATCCCGCATCTCCAGTGCATGAGCCCTTCTGGCTCCCTGTATAGATCCA 2190  
Db 200 oGlyAlaLysGlyGluValGlyProAlaGlySerProGly----- 213  
QY 2189 GACTGAACCCCTTGAAGGCCTCCAGTGCAGGAGCCCTAGAGACTGGGAGAGGAG 2130  
Db 214 -----SerAsnGlyAlaPro-----GlyGlnAr 221  
QY 2129 AGGAGGCCCGCCAGCCAGCTGTGACGTACGCACCTACGACGACGAGGTGGCAGCAG 2070  
Db 221 gGlyGluProGlyProGlnGly-----HisAlaGlyAlaGlnGlyProProGly 237  
QY 2069 AGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGCCAGCCCGCATGGGC 2010  
Db 237 yProProGlyLeuAsnGlySer-----ProGlyGly-- 247  
QY 2009 TAACAGGAGCGGAGCTGGGACCCAGTGCAGGAGGCCCTCCACCCCAATGTGCTGAAG 1950  
Db 248 -----LysGlyGluMetGlyProAla--GlyLeuProGlyAlaPro----- 260  
QY 1949 TTTTCTAGCTGAGTATTGGCCAAAGTCGTCTTGTCAATACTACTGTGTAGCAAGT 1890  
Db 261 -----GlyLeuMetGlyAlaArgGly-----ProProGlyProAla 272  
QY 1889 AAATGGGACAGACCCAGGCTGGCCAGACACATATAGGAGTGCAGTGCCTGA 1830  
Db 273 GlyAlaAsnGlyAlaProGlyLeuArg-----GlyGly 283  
QY 1829 GCTGGACAATGGACCCCAATAACAGGATGGGCCACCTGGGACAGCAGGAAGCAC--- 1773  
Db 284 AlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyProArgGlyGluArgGly 303  
QY 1772 -----TATCCAGGATGGCAGGT-----CCA 1752  
Db 304 GluAlaGlyLeuProGlyValProGlyAlaLysGlyGluAspGlyLysAspGlySerPro 323  
QY 1751 GGCAGATGCCCGCCGCGAACACCTGCTGCTGGTGGCTCACCACACACACGTA 1692  
Db 324 Gly-AspProGlyAlaAsnGlyLeuProGlyAlaAlaGly-----GluAr 338  
QY 1691 CGGAGACATCACAGGCGCCCGCA-----CAGCGCGGTG 1653  
Db 338 gGlyAlaLeuGlySerArgGlyProAlaGlyProAsnGlyLeuProGlyLysGlyP 358  
QY 1652 GAGTGGGAGCAGGCGCCACTGCCTCCAGCACCCAGTGTCCATTAGG---AAGGGAGCTC 1596  
Db 358 roAlaGlyGluArgGly---AlaProGlyProAlaGlyProArgGlyAlaAlaGlyGluP 377  
QY 1595 CAGCTTAGG-----CTTGGCAGGAAGTGTGTATCATAGCTGTCTCTACTCTAGCAC 1542  
Db 377 roGlyArgGlyValProGly-GlyProGlyMetArgGlyMetPro----- 392  
QY 1541 CTCCAGTGTCCCTCGTATTGGGCAGGAACACCTGCTTCTCCCGTGTGTAGAGGAGG 1482  
Db 393 -----GlySerProGlyProGlySerAspGlyLys 403  
QY 1481 CCAGTGTGTAGGCGCAGGCTGCAGGCTGAGAAGTGAACCCGSGTGAAGGCGGCTGAAG 1422  
Db 404 Pro---GlyProProGlySer-GlnGlyGluSerGlyArgProGlyProProGlyProse 422  
QY 1421 CTGTACACGCGGCACACTGTGGGACAGGATGTGGCCAGCGGCGGACGACGGAAGAGCTG 1362  
Db 422 r-GlyProArgGlyGlnProGlyVal---MetGlyPheProGlyProLysGlyAsn--- 439  
QY 1361 CCACACTGGCCAAATAGACTGTCTGAGTGGCGAATCGCTGCACACGCGGTCCATGACCA 1302

Db 440 AspGlyAlaProGlyLysAsnGlyGlu-----ArgGlyGlyProGlyGlyProGlyPro 457  
QY 1301 GAGAGAAGACACAGGAGATGGCCACTGCAGGA-----ACAGCC 1263  
Db 458 GlnGlyProProGlyLysAsnGlyGluThrGlyProGlnGlyProProGlyProThrGly 477  
QY 1262 CAGGCTGCCCATCCGACGCTTCATCATAGTGTCTCCGGCTCGGTGCCGCTCAG 1203  
Db 478 ProGlyGlyAspLys-----GlyAspThrGlyProArgGlyProGlnGly 492  
QY 1202 CTCTGGCCAGCCCTGTGTACA----- 1182  
Db 493 LeuGlnGlyLeuProGlyThrGlyGlyProProGlyGluAsnGluLysProGlyGluPro 512  
QY 1181 -----GCC 1179  
Db 513 GlyProLysGlyGluAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAla 532  
QY 1178 CTTGCGCCACGAATCCGTGTAACAGCGTGAAGTGCATGAGTGCATCCAGCTGCACA 1119  
Db 533 ProGlyGluArgGlyPro-----ProGlyLeuAlaGly 543  
QY 1118 GCTCAGCCAGAGAGCCGCGCAGGTCGCGGGGCATCGCGCAGCAGCAGCTGGTCAGCC 1059  
Db 544 AlaProGlyLeuArgGlyGlyAlaGlyProProGlyProGlyGlyLysGlyAlaAla 563  
QY 1058 GGGAGAGCAGGCGCCAGGTTCC----- 1035  
Db 564 GlyProGlyProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGlu 583  
QY 1034 -----GAAAGCCAGCGCGCCGCG 1014  
Db 584 ArgGlyGlyLeuGlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGly 603  
QY 1013 ATGGACAGCAGTGGGGCGACA-----AGGAGGGGCCAGACGCCCTTCTGCTGCTCGG 960  
Db 604 AlaAspGlyValProGlyLysAspGlyProArgGlyProThr----- 617  
QY 959 TGGGGCCAGCGCTGCTCTCAGCCAGCAGCAGTGTGCTGCTAGCAGGTGAGGAGCA 900  
Db 618 ---GlyProLeuGlyProGlyProAla-----GlyGln 628  
QY 899 TGAGGTGAGCAGGCCAAGAGGCACCTCTCTGGTGGCCAGGT---AGGGGGCCAGG 843  
Db 629 ProGlyAspLysGlyGluGlyAla---ProGlyLeuProGlyLeuAlaGlyPro--- 646  
QY 842 CACTGTGTCCAGTCAATGGCAGCAGGAGTAGCCAGGAGGCGCCCAAGCTGATCA 783  
Db 647 -----ArgGlySerProGlyGlu----- 652  
QY 782 TGAAGCATAGACAGTAGCTGCGACAGTGT---CCGGGT---CCCGGAAGAGGT 729  
Db 653 ---ArgGlyGluThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsn 671  
QY 728 CAGAGAGCAGGCCCTCCAGTGGAGTGAACACACCTGGCCACAGAGTCCAGCAGCCCA 669  
Db 672 GlyGluProGlyGlyLysGlyGluArgGlyAlaProGlyGlyLysGlyGluGlyPro 691  
QY 668 CCGCCAGGATGAGCAGTGCACCTCCAGG-----GCCTGGATCCGGGCGCACAGCA 618  
Db 692 ---ProGlyValAlaValProGlyGlySerGlyProAlaGlyProProGlyProGln 710  
QY 617 GCCCTGTAGCCAGCGCGCTTGGGATGAGAAAGCTCAGAGGTGAGGAGTGCACAGGACA 558  
Db 711 GlyValLysGly-----GluArgGlySerProGlyGlyPro----- 722  
QY 557 GTGCCAGATGAAGGCGCGCGCCCATAGCTCCAGCGTGTGCTCAGTGGCTGAGC 498  
Db 723 -----GlyAlaAlaGlyPhe----- 727  
QY 497 CTAGGAGCGGGACAGACAGCAGCCAGCTGGACCAATGCCAGCAGCACTGATGTTCA 438  
Db 438 ----- 438



Db 175 ----- 175  
QY 2423 GGAGGAGTGTTCATCTCTAGAACTCCATGCAAGAGCTACATTAAGAGAGCTGCAGGT 2364  
Db 175 ----- 175  
QY 2363 TAAGGGCTTAGA-----GATGGGAACACAGGTGACTGAGTTTATTTCAGCTCCCAAAA 2310  
Db 176 -AspGlyAlaArgGlySerAspGlyGlnProGly- 189  
QY 2309 CCTTCTCTAGGTGTCTCACTCACTAGGAGGCTAGCTGTTAACTCCCTGAGCCCTGGTAATCC 2250  
Db 189 yProPro---GlyThrAlaGlyPhePro- 200  
QY 2249 ACCTGCAGAGTCCCGCATTCAGTGCATGAGGCCCTTCTGCCTCCCTGTTAATAGTCCA 2190  
Db 200 cGlyAlaLysGlyGluValGlyProAlaGlySerProGly- 213  
QY 2189 GACTCAAAACCCCTTGGAGGCTCCAGTCAAGGAGGAGGAGGAG 2130  
Db 214 -----SerAsnGlyAlaPro- 221  
QY 2129 AGGGAGGCCCCAGCCCGCTGTGCAGCTACGACCTCAGCAGCAGCAGGAGTGCAGCAG 2070  
Db 221 gGlyGluProGlyProGlyProGly- 237  
QY 2069 AGAGCCACATTACTTTGGCAGCAGACAGAACTGGGGCCAGCCCGGAGGAGGAG 2010  
Db 237 yProGlyIleAsnGlySer- 247  
QY 2009 TAACAGGAGCGGGAGCTGGGACCCAGTGAGGAGCCCTCCACCCCAATGTGCTGGAAG 1950  
Db 248 -----LysGlyGluMetGlyProAla--GlyIleProGlyAlaPro- 260  
QY 1949 TTTTCTACGCTGAGTATTGGCCAAAGTCCCTCTTGTCAAACTACTACCTGTGTAGCAAGT 1890  
Db 261 -----GlyLeuMetGlyAlaArgGly- 272  
QY 1889 AAATGGCGACAGACCCAGCCCTCGCGGAGACACCATATAGCAGTGCAGACTGGCTGA 1830  
Db 273 GlyAlaAsnGlyAlaProGlyLeuArg- 283  
QY 1829 GCTGACAAATGGAGCCATAACAGGAGTGGGGCCACCTGGGACAGCAGAGAGGCAC- 1773  
Db 284 AlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyProArgGlyGluArgGly 303  
QY 1772 -----TATCCAGATGGCCAGGT- 1752  
Db 304 GluAlaGlyIleProGlyValProGlyAlaLysGlyGluAspGlyLysAspGlySerPro 323  
QY 1751 GGCAGATGCCCGCGCCGGAACCCCTCGCTCGGTGGCTCAGCCACCCACACACGTA 1692  
Db 324 Gly-AspProGlyAlaAsnGlyLeuProGlyAlaAlaGly- 338  
QY 1691 CGGAGACATCAGCAGGAGGCCCGCA-----GAGCGGGGTG 1653  
Db 338 gGlyAlaLeuGlySerArgGlyProAlaGlyProAsnGlyIleProGlyGluLys-GlyP 358  
QY 1652 GAGGTGGAGCAGCCACTGCCTCCAGCAGCCACCTGTCATAGG- 1596  
Db 358 roAlaGlyGluArgGly- 377  
QY 1595 CAGGCTTAGG-----CCTGGCAGGAAGTGGTATCAGGCTGCTCCTCACTGTAGCAC 1542  
Db 377 roGlyArgAspGlyValProGly-GlyProGlyMetArgGlyMetPro- 392  
QY 1541 CTCACGTGTCCTCGGTATTGGGAGGAGAACACCTGCTTCCCGGTGGTAGAGGAGG 1482  
Db 393 -----GlySerProGlyGlyProGlySerAspGlyLys 403  
QY 1481 CCACTGTAGGCGCAGATCTCAGGCTCAGNAGGTGAACCGGTGAGCGGCTGAAG 1422  
Db 175 ----- 175  
QY 404 Pro---GlyProProGlySer-GlnGlyGluSerGlyArgProGlyProProGlyProSe 422  
1421 CTGTCAACGAGCCACACTGTGGGACAGCATGTGGCACCGGACAGCAGGAAAGCTG 1362  
Db 422 r--GlyProArgGlyGlnProGlyVal---MetGlyPheProGlyProGlyLysAsn--- 439  
1361 CCACACTGGCCAATAGACTGCTCGAGTGCCTGCAATCGCTCCACCGCCGCTCATGACCA 1302  
Db 440 AspGlyAlaProGlyLysAsnGlyGlu-----ArgGlyGlyProGlyGlyProGlyPro 457  
QY 1301 GAGAGAGACAGGAGGAGTGGCCACTGCAGGA-----ACAGCC 1263  
Db 458 GlnGlyProGlyLysAsnGlyGluThrGlyProGlnGlyProProGlyProGly 477  
1262 CAGGCTGCCATCCGAACGCTTCATAGTGTCTCGGGCTCGGTGCCGCGCTCAG 1203  
Db 478 ProGlyGlyAspLys-----GlyAspThrGlyProArgGlyProGlnGly 492  
QY 1202 CTCTGGCAGCCCTGGTACA----- 1182  
Db 493 LeuGlnGlyLeuProGlyThrGlyGlyProProGlyGluAsnGluLysProGlyGluPro 512  
1181 -----GCC 1179  
513 GlyProLysGlyGluAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAla 532  
1178 CCTCGCCACGAAATCGTGTAAACACAGCGTGAAGTTCATAGTGCATCCAGCTGCACA 1119  
Db 533 ProGlyGluArgGlyPro-----ProGlyLeuAlaGly 543  
QY 1118 GCTCAGCAGACAGAGCCCGCAGGGTGGGGCATCGCGGACAGCAGCAGCTGTGTCAGCC 1059  
Db 544 AlaProGlyLeuArgGlyGlyAlaGlyProProGlyProGlyGlyLysGlyAlaAla 563  
QY 1058 GGGAGACAGCGCCCGCTTC----- 1035  
Db 564 GlyProProGlyProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGlu 583  
1034 -----GGAAGCCACCGGCGCCG 1014  
584 ArgGlyGlyLeuGlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGly 603  
QY 1013 ATGCACAGCAGTGGGGCACA-----AGGAGGGCGCCGACAGCCCTCTCTGCTGCTCGG 960  
Db 604 AlaAspGlyValProGlyLysAspGlyProArgGlyProThr----- 617  
QY 959 TGGGGCCAGCGCTGCTCTCAGCCACAGCAGTGTGCTGTACGAGTGCAGGAAGA 900  
Db 618 --GlyProIleGlyProProGlyProAla-----GlyGln 628  
QY 899 TGAGGGTACAGAGCCCAAGAGCAGCTCTCTCTGGTGGCCAGGT---AGGGGCCAGGG 843  
Db 629 ProGlyAspLysGlyGlyGlyAla---ProGlyLeuProGlyIleAlaGlyPro--- 646  
842 CACTGGTCTCCAGTCAATGCGCAGGAGGTAGCCAGCAGCAGCCCAAGACTGATCA 783  
647 -----ArgGlySerProGlyGlu----- 652  
QY 782 TGAAGGCATAGACAGTAGCTGGCCAGCTGGT---CCGGGT---CCCGGAAGAGGT 729  
Db 653 --ArgGlyGluThrGlyProGlyProAlaGlyPheProGlyAlaProGlyGlnAsn 671  
QY 728 CAGAGCAGGCGCTCCAGTGCAGTGAAGCAGCAGCTGGCCACAGAAAGTCCAGCAGCCCA 669  
Db 672 GlyGluProGlyLysGlyGlyAlaProGlyGluLysGlyGlyGlyPro 691  
QY 668 CGCCAGGATCAGCAGTCCAGCTCCAGG-----GCCTGGGATCCGGGACAGCA 618  
Db 692 ---ProGlyValAlaValProGlyGlySerGlyProAlaGlyProProGlyProGln 710  
QY 617 GCGCTGTAGCAGCGCCGCTGGTGGATGAGAAAGAGCTCAGCAGGATGCCCAAGGACA 558  
711 GlyValLysGly-----GluArgGlySerProGlyGlyPro----- 722

```

QY 1409 GCCGTGTGACAGCTTCAGCGCCCTCAGCGGTTACCTTCTCAGCCCTGACAGATCTG 1468
Db 481 yPro-----GlnGlyLeuProGlySerProGlyAlaProGly-ThrProGlyPro----- 497
QY 1469 CCCTACACACTGGCTCCCTCTACACCGGAGAACAGGTGTTCT----- 1515
Db 498 -----GlnGlyLeuProGlySerProGlyAlaProGly-ThrProGlyGluGlnG 515
QY 1516 -----GCCAAATACCGAGGAGACACTGGAGGTCTAGCAGTGGAGC 1558
Db 515 lnHisLeuGlyGlyAla-LysGlnAlaGlyAspValGlySerProGlyAlaProGly 534
QY 1559 AGCTGATGACAGCTTCCTGCGCAGGCCCT-----AACCTGGAGCTCCCTCCCTAAT 1612
Db 535 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGly-ThrProGlyProGln 554
QY 1613 GGACAGTGGGTGCTGAGGAGCAGTGGCTGCTCCACCTCCACCGCGCTGCGGGGCC 1672
Db 555 GlyLeuProGlySerProGlyAlaProGly-ThrProGlyProGlnGlyLeuProGlySer 574
QY 1673 -----TCT 1675
Db 575 ProGlyAlaProGlyThrProGlyGluGlnGlnHisLeuGlyGlyAlaLysGln 594
QY 1676 GCCTGTGATGCTCTCCGTACGTGTGGTGGTGAGCCACCGAGCGGAGGTGTTCCG 1735
Db 595 AlaGlyAsp-----ValGly-SerProGlyAlaPro-GlyThrProG 608
QY 1736 GCGCGGGGATCTGCCTGGA-----CCTGCGCATCTCGA---TAGTGCCTTC 1780
Db 608 lProGlnGlyLeuProGlySerProGlyAlaProGly-ThrProGlyProGlnGlyLeuP 628
QY 1781 CTGCTGCTCCAGTGGCCCATCCCTGTTTATGGCTCCATGTCACGCTCAGCAGTCT 1840
Db 628 roGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro----- 646
QY 1841 GTCATGCTATATGTGTCTGCCGAGCGCTGGTCTGGTGC-----CATTTACTTT 1894
Db 647 -----GlyAlaProGlyThrProGlyGluGlnGlnHisLeu----- 660
QY 1895 GCTACACAGTAGTATTTCACAGAGCAGCTGGCCAAATACTCAGCCTAGAAACTTCC 1954
Db 661 -----GlyGly-----AlaLysGlnAlaGlyAspValGlySerProGlyAlaProG 676
QY 1955 AGCACATTGGGTGGAGGCGCTGCTCAGTGGGTCCAGCTCCGCTCCTGTAGCCCC 2014
Db 676 lThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 695
QY 2015 ATGGGCTGCGGGCTGGCGG 2035
Db 696 GlnGlyLeuProGlySerPro 702

RESULT 25
US-08-963-825-21
; Sequence 21, Application US/08963825
; Patent No. 6110689
; GENERAL INFORMATION:
; APPLICANT: Qvist, Per
; APPLICANT: Bonde, Martin
; TITLE OF INVENTION: A Method for Assaying Collagen Fragments
; TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the
; TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
; TITLE OF INVENTION: Disorders Associated with the Metabolism of
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963,825
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,319
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda C
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 4305/08701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (III)
; US-08-963-825-21

Alignment Scores:
Pred. No.: 6,94e-20 Length: 1078
Score: 364.50 Matches: 290
Percent Similarity: 30.79% Conservative: 57
Best Local Similarity: 25.73% Mismatches: 354
Query Match: 5.86% Indels: 426
DB: 3 Gaps: 62

US-09-759-143-110 (1-3410) x US-08-963-825-21 (1-1078)
QY 2939 GGAAGTGGGGGCAACAGGCTGGGCGAAGAGAGGGGTGGTTAGGAGCGCTTGAGA 2880
Db 48 GlySerProGlyThrGlnGlyProGlyGluProGlyGlnAlaGly---ProSerGly 66
QY 2879 CCTGAAGCCCGCCCTCTACCTTCTTCAACACCCCTAACCTGGGTAACAGCATTTGGAA 2820
Db 67 ProProGlyProGly----- 72
QY 2819 TTATCATTTGGGATGAGTAGAATTTCCAAGTCTCTGGTTAGCATTTTGGGGGGCCAGA 2760
Db 73 ---AlaIleGlyProSerGlyProAlaGlyLysAspGlyLysSerGlyArgProGlyArg 91
QY 2759 CCCAGGAGAGAAGATTTCTGGCAATGATCAGCCCAATGACACGCTATCTCAGGGGACCT 2700
Db 92 ProGlyAspArgGlyLeu-----ProGlyProGlyIleLysGlyPro 106
QY 2699 GATTGTTGGGATCCCCCCTTACCAATATTAGACACCAACACAGA----- 2651
Db 107 Ala-----GlyIle-ProGlyPheProGlyMetLys---GlyHisArgGlyPheAspG 123
QY 2650 -----AAAGCTAGCAATGATCCCTTCTTCTTAAATAAATAAGTTAA 2604
Db 123 yArgAsnGlyGluLysGlyThrGlyAlaProGly-----LeuLys 137
QY 2603 ATATTTAAATGCTGTCTCTGTGATGGCAAGAGCAACAGCCACCATCTGAT 2544
Db 137 SGlyGluAsnGlyLeuProGlyGluAsnGlyAlaProGlyPrometGlyProArgGlyAl 157
QY 2543 AAAAGTAAGAGGGGGTGGATCAGCAAAAAGACAGTGTGTGGTGGGAGGACCTGGT 2484
Db 157 aProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaArgGlyAsn----- 175
QY 2483 TCTTGTGTGTTGCCCTCAGGACTCTTCCCTACAAAATAGTCATATGTTCAATCCCAT 2424

```





QY 231 AGCAGTTCTGAGTGCCTGAACGGCCCTGAGCCCTACCCGCTGCGCCCACTATGCT-- 288  
Db 111 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro-GlnGlyLe 130  
QY 289 -----CCAGAGCTGTGGTGAAGCG--CCTGCTGCGGACCGAAGCCAGC 335  
Db 130 uProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProG 150  
QY 336 TCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTGGCGGAGGACATCACT 395  
Db 150 yAlaProGlyThrProGlyGlu-----GlyGlnGlnHisHisLe 163  
QY 396 ATGTGCGCCCTGCTGCTGCTGGAAGTGGGGTAGAGGAGAGTTTCATGACCATGCTGG 455  
Db 163 uGlyGlyAlaLysGlnAlaGlyAspValGly----- 173  
QY 456 GCATTGTGTCAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 515  
Db 174 -----SerProGlyAlaProGly--ThrProGlyProGln----- 184  
QY 516 GCGTGGACCTATGCGCGCGCGCGGCTTCATCTGGGACATGCTTGGGACATCTGCTG 575  
Db 185 -----GlyLeuProGlySerProGly 191  
QY 576 TGACCTP---CTTCTCATCCCAAGGCGGCTGGCTAGCAGGGCTGCT-----GTGCC 626  
Db 191 yAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 211  
QY 627 CGATGCCAG-----GCCCTGGAGCT---GGCAGTGTCTCATCTCTGGGGTGG 671  
Db 211 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-----ProGlyGluGl 229  
QY 672 G-----GCTGCTGGACTTCTGCGCCAGGTGCTTCACTCACTGGAGGCC 719  
Db 229 yGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaPr 249  
QY 720 TGCTCTGTGACCTTCTCCGGGAGCC-----GGACCACTGTC 755  
Db 249 oGly-----ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 266  
QY 756 GCAGAGCTACTCTGCTATGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815  
Db 266 rProGly-----ProGlnGlyLeuProGlySerProGly 277  
QY 816 TGCTGCTCATTTGACTGGACACAGTGCCTGCCCC-----CTACCTGGGACCCAGG 869  
Db 277 yAla-----ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 291  
QY 870 AGGAGTGCCTTTTGGCCCTGCTCACCTCATCTTCTCATCTGCTGCTGCTGCTGCTGCTG 929  
Db 292 -GlyAlaProGlyThrProGlyGluGlnGlnHisHisLeuGlyGlyAlaLysGlnAl 311  
QY 930 TGCTGCTGAGGAGGAGCGCTGGGCCCCCAGCCAGCCAGCAGAGGCTGCTGCGCCCTCT 989  
Db 311 aGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly 331  
QY 990 CTTTGTGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1046  
Db 331 rProGlyAlaPro-----GlyThrProGlnGlyLeuProGlySerProGlyAla 349  
QY 1047 CCTGCTTCCCGGCTGACAGCTGTGCTGCGCATGCGCCGACCTGCTGCGCCGCTCT 1106  
Db 349 aProGlyThrPro--GlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 367  
QY 1107 TCGTGGCTGAGCTGTGCAG-----CTGATGGCACTATGACCTTACGCTGTT----- 1155  
Db 367 rProGlyGluGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGly 387  
QY 1156 -----TTACACGGATTTCTGGGCGAGGCGCTGACAGGCGCTGCGCCAGCTGAGC 1208  
Db 387 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla--pr 406  
QY 1209 CGGGCAGGAGCCCGGAGACACTATGATGAAGGCTTCCGATGGCAGCCTGGGGCT-- 1266

Db 406 oGlyThrProGlyProGlnGlyLeu-----ProGlySerProGlyAlaPr 421  
QY 1267 -----GTTCTCTGAGTGCCTCATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1322  
Db 421 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 441  
QY 1323 AGCGATTCCGGAC-----TCGAGCAGTCTATTGGCCAGTGTGGCAGCTT-----TC 1369  
Db 441 uGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAl 461  
QY 1370 CCTGTGGCTGGCTGCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1408  
Db 461 aProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGl 481  
QY 1409 GCGTGGTGGACAGTTCAGCCGCTCACCAGGTTTCACTTCTCAGCCCTCAGATCCTG 1468  
Db 481 yPro-----GlnGlyLeuProGlySerProGlyAlaProGly--ThrProGlyPro 497  
QY 1469 CCCTACACACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515  
Db 498 -----GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlnG 515  
QY 1516 -----GCCCAATACGAGGAGGACACTGGAGTGTGCTAGCAGTGGAGC 1558  
Db 515 lnHisHisLeuGlyGlyAla-LysGlnAlaGlyAspValGlySerProGlyAlaProGly 534  
QY 1559 AGCTGTGATGACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1612  
Db 535 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGln 554  
QY 1613 GGACAGTGGTGGTGGAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1672  
Db 555 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 574  
QY 1673 -----TCT 1675  
Db 575 ProGlyAlaProGlyThrProGlyGluGlnGlnHisHisLeuGlyGlyAlaLysGln 594  
QY 1676 GCTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1735  
Db 595 AlaGlyAsp-----ValGly-SerProGlyAlaPro-GlyThrProG 608  
QY 1736 GCGCGGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1780  
Db 608 lyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuP 628  
QY 1781 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1840  
Db 628 roGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 646  
QY 1841 GTCAGTGCCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1894  
Db 647 -----GlyAlaProGlyThrProGlyGluGlnGlnHisHisLeu 660  
QY 1895 GCTACAGAGTGTATTTGACAGAGCGACTTGGCCAAATACTCAGCGTGAAGAACTTCC 1954  
Db 661 -----GlyGly-----AlaLysGlnAlaGlyAspValGlySerProGlyAlaProG 676  
QY 1955 AGCATTTGGGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2014  
Db 676 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrProGlyPro 695  
QY 2015 ATGGGCTGCGGGCTGCGCG 2035  
Db 696 GlnGlyLeuProGlySerPro 702

## RESULT 24

US-08-397-633A-26

: Sequence 26, Application US/08397633A

: Patent No. 5773577

: GENERAL INFORMATION:

: APPLICANT: Cappello, Joseph





2045 QY -----CAGAAACTGGCGCCAGCCGCGCAGCCGATGGGCTAACAGGAG 2001  
100 Db GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 119  
2000 QY CGGGAGCTGGAGCCAGTAGGAGCGCCCTCCACCCATGCTGCTGGAAGTTTCTACG 1941  
120 Db GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly 132  
1940 QY CTGAGTATTGGCCAAGTCGCTCTGTGCAATACCTACCTGTGTAGCAAAATAATGSCGA 1881  
133 Db -----GlyProGlnGlyLeuProGlnGlyLeuProGlnGlyLeuProGlnGlyLeuPro 138  
1880 QY CCAGACCCAGCCCTGCGGACAGACACCATATAGGCAGTAGCAGACTGGCTAGCTGGACAA 1821  
139 Db GlySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGly--- 157  
1820 QY TGGAGCCATAACAGGGATGGCGCCACCTCGGACAGCAGCAAGGAC----- 1773  
158 Db ---SerPro-----GlyAlaProGlyThrProGlyProGlnGlyLeuProGly 172  
1772 QY -----TATCCAGATGGCAGGAGTCCAGGCAGATGCCCGGC-----CCG 1734  
173 Db SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 192  
1733 QY GAACACCCCTGGCTCGGTG-----GGCTACCCACCACACACCTAGC----- 1690  
193 Db GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGly 212  
1689 QY -----GAGACATCACAGGACGCGCCGCGCAGAGCGCGGTGGAGTGGGACAGGCCA 1636  
213 Db AlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGly 232  
1635 QY CTGCTCCAGCACCCAGCTGTCCATTAGGAAGGAGGCTCCAGGCTTAGGGCTTGGCAGG 1576  
233 Db LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu---ProGly-Se 251  
1575 QY AAGCTGTATCAGCTGTCTCTACTGTAGCAGCTCCAGTGTCCCTCGGTATTGGC 1516  
251 Db rProGlyAlaProGlyThrProGly-----ProGlnGlyLeu----- 263  
1515 QY AGGAACACCTCTCTCCCGTGTAGAGGAGGCCA-----GTGTGTAGGCGAGGATC 1462  
264 Db -----ProGlySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAs 281  
1461 QY TCCAGGCTGAGAAGTGAACCCGCTGAGCGCGCTGAAGCTGTACACCGCCGACACTG 1402  
281 Db pValGly-----SerProGlyAlaProGly 289  
1401 QY TGGNACAGGCATGTGGCACCGGACGACAGGGAAGCTGCCACACTGGGCAATAGACT 1342  
289 Db yThrProGly-----ProGlnGly---LeuProGlySerPro----- 300  
1341 QY GCTCAGTGGCGAATTCGCTGACACCGCGGTCCATGACAGAGAGACAGGAGATG 1282  
301 Db -----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPr 315  
1281 QY CGCACTGCAGGAACA-----GCCCGAGGCTG 1255  
315 Db oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 335  
1254 QY CCCA---TCCGAAGCCTTATCATATAGTCTCCGGCTCGGTGCCCGCTAGCTCTG 1198  
335 Db rProGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGly 355  
1197 QY GGCAGC-----CCTGTATACGCCCCTCGCCACAGAAATCCGTTAAACAGCGTGAAG 1144  
355 Db yProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGly----- 373  
1143 QY GTCATGAGTGCATCCAGC-----TGCACAGCTCAGCCACAGAGCGCGC 1096  
374 Db -----LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly 391  
1095 QY AGGTGCGGGGCATGCGGCACACAGCTGTGTGCA-----GCCGGGAGCAGGCGGCC 1042

391 Db ySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerPr 411  
1041 QY AGTTTCC---GGAAAGCCAGCGCGCCGATGGACACAGCTGGGGCCACAGGAGGG 985  
411 Db oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 431  
984 QY GCGACAGCCCTTCTGCTGCTGGTGGGGCCCA---GGCTGCTCCCTCCACGCCACGAC 928  
431 Db rPro-----GlyProGlnGlyLeuProGlySerPro----- 441  
927 QY AGTGTGGCTGTACGCAGGTGAGGAAGATGAGGGTGAAGGAGCAAGAGGACCTCT-- 870  
442 Db -----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlyProGlnGlyLeuProGln 454  
869 QY ---CCTGGTGGTCCAGGTAGGGGCGCAGGCGACTGGTCTCCAGCTCAATGCGCAGCAG 814  
454 Db ySerProGlyAlaProGlyThr---ProGly-----GlyAlaLysGlnAlaGln 469  
813 QY A-----GGTAGCCCGCAGGAGCCCAAGACTGATCATGAAGCATAGACAGTAGGCC 760  
469 Db yAspValGlySerProGlyAlaPro-----GlyThrPr 480  
759 QY TGGCGACAGTGT---CCGGGT---CCCGAAGAGGTACAGAGCAGGAGGCTCCAGTGA 706  
480 Db oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 500  
705 QY GTGAACACACCTGGCCACAGAGTCCAGCAGCCCGCCAGGATGAGCAGTGCACG 646  
500 Db uProGlySerProGly-----AlaProGlyThrProGlyProGlnGlyLeuProGlySe 518  
645 QY TCCAGGGCTGGGATCCGGGCACAGCAGCCCTGTAGCAGCCGCGCTTGGGATGAGA 586  
518 Db rProGlyAla-----ProGlyThrProGly-----GlyAlaLysGlnAlaGly---As 533  
585 QY AAGAGCTCAGCAGGATGCCAAGGACAGTG---CCCATATGAGGCGCGGCGGCCA 529  
533 Db pValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly----- 550  
528 QY TAGCGCTCAGCCAGTGGTCACTGGCTGAGCCCTAGGAGCGGACACAGCAGGCGCCAGC 469  
551 Db -----SerProGlyAlaProGlyThrProGlyPro----- 560  
468 QY ACTGACCAATGCCAGCAGCACCATTGTCATGAATCTCTCTACCCCTCCAGCAGCAG 409  
561 Db -----GlnGlyLeuProGlySerProGlyAl 569  
408 QY AGAGCGGCACATAGTGATGCTCGCGCCCAACACACCTCCAGGCCAAGGTTAGCAGG 349  
569 Db a-----ProGlyThrProGlyProGlnGlyLeuProGln 580  
348 QY TTGACCAAGAGCTGGGCTTCCGGTGCCGCA-----GCAGGC-----GGCTCACC 301  
580 Db ySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerPr 600  
300 QY CACAGCCTCTGGACCATAGTGGGCCAGGCGG---TAGGGCTCAGGGGCGGTTACAGCA 244  
600 Db oGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyT 620  
243 QY CTCGCAACTGCTCTGCTCGGCTCTCCAGAACTGCGGCTCTCTCTCTCTCTCTCTCTG 184  
620 Db hrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly----- 637  
183 QY GCCAACTGCTTAGGAATCAGCAGGCGCCATTTCCTGCCAGCCCTTGTGTCGCGGCCAG 124  
638 Db -----ProGlnGlyLeuProGlySer-----ProGlyAlaProGlyThrP 651  
123 QY CTCTCAGCCCATGCTCAACACCTGCTGTGTGGGACACTCAGTGGGACAGCTCTCAT 64  
651 Db rogly-----GlyAlaLysGlnAlaGlyAspValGly 662  
63 QY CA-----CTCAGATCTGCCCA 46  
|| ::::|

[illegible]

```

RESULT 21
US-08-397-633A-36
; Sequence 36, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
; TITLE OF INVENTION: OF ENZYMA TIC CROSS-LINKING
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,633A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 682 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-397-633A-36

Alignment Scores:
Pred.No.:      2.04e-20          Length:      682
Score:         370.50           Matches:     249
Percent Similarity: 35.26%       Conservative: 44
Best Local Similarity: 29.96%    Mismatches:  294
Query Match:    5.95%           Indels:      246
DB:             1              Gaps:        56

US-09-759-143-110 (1-3410) x US-08-397-633A-36 (1-682)
Qy   2314 AAAAACCCCTCTCTAGGTGTCCTCAACTAGGAGGCTAGTCTTAACCCCTGAGCCTGGGT 2255
      :::::|||||
Db   12 GluaSnPro-----GlyValThrGlnLeuasn-----ArgLeuAlaAlaHis 25
      :::::|||||

Qy   2254 AATCCACCTGCAGAGTCCCGCATTCACGTCCAGTGCGATGGAGCCCTTCT- 2210
      ||||| ||||| ||||| :::::
Db   26 ProPheAlaSerAspProMetGlyAlaProGlyThrProGlyProGlnGlyLeuPro 45
      ||||| ||||| ||||| :::::

Qy   2209 GGCTCCCTGTATAAGTCAGACTCAAACCCCTTGGAAAGCCCTCCAGTCAGGCAGCCCT 2150
      ||||| ||||| ||||| :::::
Db   46 GlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeu---ProGlySerPro 63
      ||||| ||||| ||||| :::::

Qy   2149 ACAGACTGGGGAGAGAGAGAGGGAGGCCGCCAGCCAGCTGTGCAGCTACGCACCCTCA 2090
      ||||| ||||| ||||| :::::
Db   64 -----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly 79
      ||||| ||||| ||||| :::::

Qy   2089 GCA-----GCACAGSGTGGCAGCAGAGCCACATTACTTTGGCACAA----- 2046
      ||||| ||||| ||||| ||||| :::::
Db   80 AlaProGlyThrProGlyGlyAlaGlyAlaGlnAlaGlyAspValGlySerProGlyAlaPro 99
      ||||| ||||| ||||| ||||| :::::

```

US-08-642-255-126  
; Sequence 126, Application US/08642255  
; Patent No. 5773249  
; GENERAL INFORMATION:  
; APPLICANT: CAPPELLO, Joseph  
; APPLICANT: FERRARI, Franco A.  
; TITLE OF INVENTION: High Molecular Weight Collagen-Like  
; TITLE OF INVENTION: Protein Polymers  
; NUMBER OF SEQUENCES: 135  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/642,255  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROWLAND, Bertram I.  
; REGISTRATION NUMBER: 20,015  
; REFERENCE/DOCKET NUMBER: A55556-3/BIR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 494-8700  
; TELEFAX: (415) 494-8771  
; TELEX: 910 277299 FHT UR  
; INFORMATION FOR SEQ ID NO: 126:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 682 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-642-255-126

Alignment Scores:  
Pred. No.: 2,04e-20 Length: 682  
Score: 370.50 Matches: 249  
Percent Similarity: 35.26% Conservative: 44  
Best Local Similarity: 29.96% Mismatches: 294  
Query Match: 5.95% Indels: 246  
DB: 1 Gaps: 56

US-09-759-143-110 (1-3410) x US-08-642-255-126 (1-682)  
QY 2314 AAAAACCCCTTCTAGGTGTCCTCACTAGGAGGCTAGCTGTTAAACCTCAGCGCTGGGT 2255  
DB 12 GluAsnPro-----GlyValThrGlnLeuAsn-----ArgLeuAlaAlaHis 25  
QY 2254 AATCCACCTGCAGAGTCCCGCATTCACAGTGCATGGAGCCCTTCT----- 2210  
DB 26 ProProPheAlaSerAspProMetGlyAlaProGlyThrProGlyProGlnGlyLeuPro 45  
QY 2209 GGCCTCCCTGTATAAGTCACAGCTGAAACCCCTTGGAGGCGCTCCAGTCAGGAGCGCCT 2150  
DB 46 GlySerProGlyAlaProGlyThrProGlyPro-----GlnGlyLeu-----ProGlySerPro 63  
QY 2149 AGAGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2090  
DB 64 -----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly 79  
QY 2089 GCA-----GCACAGGTTGGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2046  
DB 80 AlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaPro 99  
QY 2045 -----CAGAACTGGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2001

100	GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro	119
2000	CGGGAGCTGGGACCCAGTGGAGCGGCGCTCCACCCCAATGTCTGGAAGTTTCTACG	1941
120	GlnGlyLeuProGlySerProGlyAlaProGlyThrPro-----	132
1940	CTGAGTATTGGCCCAAGTCGCTTGTCAATACCTGTCTAGCAAAAGTAAATGGGCA	1881
133	-----GlyProGlnGlyLeuPro-----	138
1880	CCAGACCCAGGCTGGCGGAGCAGACCATATAGGAGTGCACACTGGCTGAGCTGGACNA	1821
139	GlySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGly	157
1820	TGGAGCCCATAAACAGGATGGCGCACCTGGAGCAGACGAGGAGGCAC-----	1773
158	-----SerPro-----GlyAlaProGlyThrProGlyProGlnGlyLeuProGly	172
1772	-----TATCCAGGATGGCGAGTCCAGCAGATGCCCGGC-----	1734
173	SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro	192
1733	GAACACCTGGCGCTCGTG-----GCTCACCACCCACACCTACG-----	1690
193	GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGly	212
1689	-----GAGACATCACAGGCGCGGAGGCGGCGGAGGCGGCGGAGGCGGAGGCGCA	1636
213	AlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGly	232
1635	CTGCCTCCAGCACCACCGTCCATAGGAGGAGGAGCTCCAGGCTTGGCGCTGCCAGG	1576
233	LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu-----ProGly-Se	251
1575	AGCTGGTCATCAGGCTGCTCCTACCTGCTAGCACCTCCAGTGCCTCGGTATTTGGGC	1516
251	rProGlyAlaProGlyThrProGly-----ProGlnGlyLeu-----	263
1515	AGGAACACCTGCTTCTCCGCGTGTAGAGGAGGAGGCGCA-----GTCTGTAGGCGAGATC	1462
264	-----ProGlySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAs	281
1461	TGCAGGCTGAGAGGTGAACCGGTGAGGCGGCGCTGAGCTGCACCGGCGCACACTG	1402
281	pValGly-----SerProGlyAlaProGly	289
1401	TGGGACAGGATGTGGCACCGCGCACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1342
289	yThrProGly-----ProGlnGly-----LeuProGlySerPro-----	300
1341	GCTCGAGTGCAGATCGCTGCACAGCGGCTGCATGACAGAGAGAGAGAGAGAGAGAGAG	1282
301	-----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro	315
1281	CGGCACTGCAGGAGCA-----	1255
315	oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh	335
1254	CCCA---TCCGAGCGCTTCATCATAGTGTCTCGGCGCTCGGTGCGCGCTCAGCTCTG	1198
335	rProGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGly	355
1197	GGCAGCG-----CCTGGTACAGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1144
355	yProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGly-----	373
1143	GTCACTAGTGCCTCCAGC-----TGCACAGCTCAGCCAGGAGAGGAGGAGGAGGAG	1096
374	-----LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly	391
1095	AGGTGCGGGGCGATGCGGCGCAGCAGCTGTGTGCA-----GCCGGGAGAGGAGGCGCC	1042



1879 CAGACCCAGGCTGCGGCAGACACCATATAGGCACTGACAGACTGGCTGAGCTGGACAAT 1820  
184 --AspProGly-----ProProGlyAlaGlnGlyProAlaGlyProGlyGlySerA 200  
1819 GGAGCCCAACAGGAGTGGGCGCCACCTGGGACACAGCAAGGACCATATCCAGGATGGC 1760  
200 rgAspPro-----GlyProProGlyAlaGlnGly-----ProAlaGlyP 213  
1759 GAGTCCAGGACAGTGCCTGGCGCCGAA-----1731  
213 roGlyGlySerArgAspProGlyProProGlyAlaHisGlyProAlaGlyProLysGlyA 233  
1730 -----CCACCTGGCTGGTGGCTACCCACACACACAGTACGAGACATCACA 1679  
233 laHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisG 253  
1678 G-----GCAGAGCCCGCCAGAGCGGGGTGGAGTGGGAGGAGG- 1639  
253 lyProAlaGlyProLysGlyAlaGlnGlyPro-----AlaGlyProGlyGlySerArgA 271  
1638 --CCACTGCTCCAGCACCCAGCTGTCCATTAGGGAAGGAGCT-----1597  
271 spProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgAspProGlyP 291  
1596 --CCAGGC-----TTAGGCGCTGGCAGGAGCTGGTCAATCAGGCTGTCTCTCAC 1551  
291 roProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgAspProGlyProPro--- 309  
1550 TGCTAGCACCTCCAGTGTCCCTCGGTATTGTCGAGGACACCTGCTCTCTCCGCTGGT 1491  
310 -----GlyAlaGlnGlyProAlaGlyProGlyGly 319  
1490 AGA-----GGAGGCGCAGTGTGTAG-----GCAGGATCTCGAGGCTGAGAAGG 1446  
320 SerArgAspProGlyProGlyAlaGlnGlyProAlaGlyProGlyGlySerArg--- 338  
1445 TGAACCGGTGAGGCGGTGAGCTGTACACAGGCGCCACACTGTGGGACAGCATGTGG 1386  
339 -----AspProGlyProProGlyAlaGlnGly----- 347  
1385 CACCGGCGACGACAGGGAAGCTGCCACACTGCCAATAGACTGCTCGAGTGCCGAATC 1326  
347 ----- 347  
1325 GCTGCACACCGCTCCATGACAGAGAGAAAGACAGGAGATGCGGCACCTGCAGGAACA 1266  
348 -----ProAlaGlyProGlyGlySerArgAspProGlyProGlyAlaGlnGlyPro 365  
1265 GCC-----CCAGGCTGCCATCCAGCGCTTCATCATAGTGT 1227  
366 AlaGlyProGlyGlySerArgAspProGlyProGlyAlaGlnGlyProAlaGlyPro 385  
1226 TCCGGGCGCTCGGTGCGCG-----GCTCAGCTCTGGCAGCAGCGCTGTGTACAGCC 1179  
386 GlyGlySerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySer 405  
1178 CCTGCGCCAGAAATCCGTGTAACACAGCTGAGTGCATGAGTGCATCCAGCTCCACA 1119  
406 ArgAspProGlyProPro----- 411  
1118 GCTCAGCCAGAGACCGCGGAGGTGGCGGATGCGGCAGCACAGCT-----GGTGCA 1062  
412 -----GlyAlaGlnGlyProAlaGlyProGlyGly 421  
1061 GCCGGGAAGCAGGCGCCAGGTTCGGAAGCCCAAGCGCGCGCATGGACAGCAGT 1002  
422 SerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGly----- 437  
1001 GGGCGCAAGAGGCGCGCACAGCCCTTCTGCTGCTGCGTGGGCGCCAGCGCTGCCT 942  
438 GlySerArgAspProGlyProPro-----GlyAlaGlnGlyProAlaGlyPro 453

RESULT 20

941 -----CCTCAGCCACCAGCAGTGTGCTGTACCCAGGTGAGGAAGATGA 897  
454 GlyGlySerArgAspProGlyProPro-----GlyAlaHis 465  
896 GGTGTAGCAGCGCCAAAGAGGCACT-----CCTCTGGTCCCGAGGT--- 855  
466 GlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyPro 485  
854 AGGGGCGCA-----GGGCACTGTGTCTCCAGTCAATGGCAGGAGGAGGTAGC 807  
486 AlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaGlnGlyProAlaGly 505  
806 CCAGGAGCGCCCAAGACTGATCATGAAGCATAGACAGAGTAGGCTGCGCAGAGTGGT 747  
506 ProGlyGlySerArgAsp-----ProGlyProProGlyAlaGlnGly 519  
746 CCG---GGTCCCGGAAGAGGTCAAGAGCAGGAGCGCTCCAGTGGAGTGAACACACAGTGGC 690  
520 ProAlaGlyProGlyGlySerArgAspProGlyPro-----ProGly 533  
689 CACAGAAGTCCAGCAGCGCCA-----CGCCAGGATGAGCAGTG 651  
534 AlaGlnGlyProAlaGlyProGlyGlySerArgAspProGlyProProGlyAlaGlnGly 553  
650 CAGCT-----CCAGGCGCTGGATCCGGC---ACAGCAGCGCTGTAGCAGCGCG 600  
554 ProAlaGlyProGlyGlySerArgAspProGlyProProGlyAlaGlnGlyProGlySerArg 573  
599 CCCTGGGATGAGAAGAG---GCTCAGCAGGATGCCAAGGACAGTGCACAGTGAAGG 543  
573 yProGlyGlySerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyG 593  
542 GCGCGCGCGCGCTAGCTGCTCCAGCGCAGTGGTCACTGGCTGAGCTAGGAGCGGGACAC 483  
593 lySerArgAspProGlyProGlyAlaGlnGlyProAlaGlyProGlyGlySerArg 612  
482 AGACCGCGCCAGCAGTGGACCAATGCCAGCA---CCATGGTCACTGAATCTCTCTCTA 426  
613 AspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgAspProGly 632  
425 CCGCCACTTCCAGCAGCAGGCGGCACATAGTGTGCTGCGCGCCCAACACACACTCCA 366  
633 Pro-----ProGlyAlaGlnGly-----ProAlaGlyProGly 643  
365 GGCAGAGGTAGCAGGTTGACAGCAAGA-----GCTGGGCTTCCGGTGGCGCA 315  
644 GlySerArgAspProGlyProGlyAlaGlnGlyProAlaGlyProGlyGlySerArg 663  
314 GCAGGCGCTCACCCAGCAGCTCTGGACCATAGTGGCGCAGCGGGGTAGGCTCAGGGGG 255  
664 AspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgAspProGly 683  
254 CCGTTCAGGCACTCCAGAACTGCTGCTGCGCTCTCCAGAACTGCGGCTCTCC 195  
683 yProProGlyAlaGln-----GlyProAlaGlyProGlyGlySerArgAs 698  
194 TCCTTGTCTGCCCAACTGCTAGGAATCAGCAGCGCCCATTTCTGCCAGCCCTTTGG 135  
698 pPro-----GlyProProGlyAlaHis-----GlyProAlaGly 709  
134 TGCGCGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGTGTGGGCACTCAGTGGG 75  
709 yProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProL 729  
74 ACAGCTCTCATCACTCAGATCTCTGGCGGAGCGCGG-----CTGTACCCCGGAGCCA 21  
729 ysGlyAlaHisGlyProAlaGlyProLysGlyAlaGlnGlyProAlaGlyProGlyGlyS 749  
20 GC 19  
749 er 749





Db	91	HisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro	110
QY	2239	TCCCGCATTCACAGTCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAGTCCACAGCTGAAACC	2180
Db	111	GlyProGlnGlyLeuProGlySerPro---GlyAlaPro-----GlyThrProGly	126
QY	2179	CCCTTGGAAGCCCTCCAGTCAGCAGCCCTAGAGACTGGGAGAGAGAGGACGCC	2120
Db	127	Pro---GlnGlyLeu---ProGlySerPro-----GlyAlaProGlyThrPro	140
QY	2119	CAGCCCCCACTGTGCAGTACCCACCTCAGCAGACAGAGGT-----GGCAGCAGAGAG	2066
Db	141	GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln	159
QY	2065	CCACATTACTTTGGCAGCAACAGAAACTGGCGCCAGCCCGCAGCCGCCCATGGGGTAAC	2006
Db	160	GlnHisHisLeuGly-----GlyAlaLys	167
QY	2005	AGNAGCGGAGCTGGGACCCAGTAGGAGCAGCCCTCCACCCCAATGTGCTGGAAGTTT	1946
Db	168	GlnAlaGlyAspValGlySerPro---GlyAlaProGlyThrPro-----	181
QY	1945	CTAGCGTGAGTATTGGCCAGTGCCTCTGTGTCNAATACTACCTGTGTAGCAAGTAAT	1886
Db	182	-----GlyProGlnGlyLeuPro-----	187
QY	1885	GGCCACAGACCAGCCCTGCGGCAGACACCATTATAGCAGTGCACAGACTGCCTGAGCTG	1826
Db	188	-----GlySerProGlyAlaProGlyThrProGlyProGln-----	199
QY	1825	GACNATGGACCCATAAACAGGATGGGGCCACCTGGGACAGCAGGAAGGCACATATCCAG	1766
Db	200	-----GlyLeuProGlySerProGlyAla-----ProG	209
QY	1765	GATGGCGAGCTCAGGCAGATGCCCGGC-----CCGGAACCCCTGGGCTC	1718
Db	209	lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro---GlyThrProGlyGlu	228
QY	1717	GTTGGGCTCACCCACCACACAGCTACGGAGACATCACAGGCA-----GAGGCCCC	1667
Db	229	Gly-----GlnGlnHisHisLeuGlyGly-AlaLysGlnAlaGlyAspValGlySerPr	246
QY	1666	GCAGAGCGGGTGGAGTGGGAGCAGGCCACTGCTCCAGACCCAGTGTCCATTAGG	1607
Db	246	oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro-----	261
QY	1606	GAAGGAGCTCCAGGCTTAGGGCTCGCAGGAAGCTGTCATCAGGCTGCTCTCCTCAGTCT	1547
Db	262	-----GlyAlaProGlyThr---ProGlyProGlnGlyLeu---ProGlySerPro-----	276
QY	1546	AGCACCTCCAGTGTCCCTCCGTATTTTGGCAGGAACACCTGCTCTCCCGGTGGTAGAG	1487
Db	277	-----GlyAlaProGlyThrProGlyProGlnG	286
QY	1486	GGAGGCCAGTGTGAGGCAGGATCTGCAGGGCTGAGAAAGTGAACCCGTTGAGGGCGGC	1427
Db	286	lyLeuPro-----GlySerProGlyAla-ProGlyThrProGlyGluGly---	300
QY	1426	TGAAGCTGTACCACGCCACACTGTGGGACAGGATGTGCCACCGGACG-----	1377
Db	301	---GlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAla	319
QY	1376	-----CCACGGAAAGCTGCCACACTGGCCCAATAGACTGCTCGAGTGC	1332
Db	320	ProGlyThrProGlyProGlnGly---LeuProGlySerPro-----	332
QY	1331	CGAATCGCTCCACCGCGGTCCATGACACAGAGAGAACACCGAGGATGGCGACTGCA	1272
Db	333	-----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro	350
QY	1271	GGAACA-----GCCCGAGCGCTGCCATCCGAA	1245

QY 1190 CTTGGTACACCCCTCGCCACGAAATCCGTGTAAACACGCGTGAAGGTTCATGAGTGCCA 1131  
Db ||||| ||| ||||| :|||  
QY 391 ProGlyThr---ProGlyProGlnGly-----LeuPro 400  
QY 1130 TCCAGCTGCACAGCTCAGCCACGACGAGCGCGCGGCGCATCGCGCAGCACA 1071  
Db ||||| ||| ||||| :|||  
QY 401 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 420  
QY 1070 GCTGGTGCAGCGGGGAAGCAGGCGCCAGGTTC---GGAAGCCAAAGCGGCGCGGC 1014  
Db ||||| ||| ||||| ||||| :|||  
QY 421 ProGlyThrProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 440  
QY 1013 AT---GGCAGCAG---TGGGGCGACAGGAGGGCGCGACGCCCTTCTGTGGC 964  
Db ||||| ||| ||||| ||||| :|||  
QY 441 GluGlyGlnGlnHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGly 460  
QY 963 TCGGTGGG-----CCGCGCTGCTCTCAGCC 934  
Db ||||| ||| ||||| ||||| :|||  
QY 461 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 480  
QY 933 ACCAGCAGTGTGCTGTACGACAGGTGAGGAGTGTGAGCAGGCGCCAAAGAGGCAC 874  
Db ||||| ||| ||||| ||||| :|||  
QY 480 OGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 500  
QY 873 TCCT-----CCTGGTGCCAGTAGGGGCCAGGCGCAGTGTGTCTCCAG 829  
Db ||||| ||| ||||| ||||| :|||  
QY 500 uProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisLeuGlyG 520  
QY 828 TCAATGCCAGCAGGA-----GGTAGCCAGCAGCCCGCCAGACATCATGAAGCCA 775  
Db ||||| ||| ||||| ||||| :|||  
QY 520 yAlaLysGlnAlaGlyAspValGlySerProGlyAlaPro----- 533  
QY 774 TAGACAGTAGTGGCTCGGCACAGTGT-----CCGGT---CCCGGAAGAGTTCAGAGAGC 721  
Db ||||| ||| ||||| ||||| :|||  
QY 534 -----GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 551  
QY 720 AGGCGCTCAGTGGAGTGAACACACTGGCCACAGAAATCCAGCAGCCCGCCAGCAGG 661  
Db ||||| ||| ||||| ||||| :|||  
QY 551 OGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrProGlyProG 569  
QY 660 ATGAGCAGTCCAGCTCCAGGGCGCTGGGATCGGCACAGCAGCCCTGTCTGCCAGCCG 601  
Db ||||| ||| ||||| ||||| :|||  
QY 569 nGlyLeuProGlySerProGlyAla-----ProGlyThrProGlyGluGlyGlnGlnH 587  
QY 600 GCCTTGGGTGATGAAGA-----GGCTCAGCAGGATGCCCAAGGACAGT 556  
Db ||||| ||| ||||| ||||| :|||  
QY 587 sHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPr 607  
QY 555 G---CCAGATGAAGCGCGCGCGGCCCATAGCTCCAGCCAGTGTCTACTGGCTGAG 499  
Db ||||| ||| ||||| ||||| :|||  
QY 607 OGlyProGlnGlyLeuProGly-----Se 615  
QY 498 CTTAGAGCGGACGACAGACAGCGCCAGCAGCTGGACCAATGCCAGCACCATTGGTCATG 439  
Db ||||| ||| ||||| ||||| :|||  
QY 615 rProGlyAlaProGlyThrProGlyPro----- 624  
QY 438 AACTTCTCTCTACCCCACTTCCAGCAGCAGAGCGGCGCACATAGTGTGTCTGGCGCC 379  
Db ||||| ||| ||||| ||||| :|||  
QY 625 ---GlnGlyLeuProGlySerProGlyAla-----Pr 634  
QY 378 AACAACACCTCAGGCGCAAGGTTAGAGCTTGACACAGCAAGAGCTGGGGTTTCCGGT--- 321  
Db ||||| ||| ||||| ||||| :|||  
QY 634 OGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyG 654  
QY 320 -----GCGCAGCAGCGCGCTCACCACAGC 295  
Db ||||| ||| ||||| ||||| :|||  
QY 654 uGlyGlnGlnHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAl 674  
QY 294 CTCTGTACCATAGTGGGCCAGCGCGG---TAGGGCTCAGGGGCGCGTTCAGGACATCCAG 238  
Db ||||| ||| ||||| ||||| :|||  
QY 674 aProGlyThr---ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProG 694

QY 237 AACTGTCTCTCTCGCTCTGCTCCAGAAAGCTGCGGCTCTCTCTCTCTCTCTCTCTCTCT 178  
Db ||||| ||| ||||| ||||| :|||  
QY 694 lyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly----- 709  
QY 177 TGCTAGGATCAGCCAGGCGCCCATTTCTGCGAGCCCTTTGGTGGCGGTCCAGCTTCTC 118  
Db ||||| ||| ||||| ||||| :|||  
QY 710 --ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGluGlyGlnG 728  
QY 117 AGCCCATGCTCAACACCTGCTGTGGGCGACCTCAGTGGGACACAGTCTCATCA--- 62  
Db ||||| ||| ||||| ||||| :|||  
QY 728 lnhHisLeuGly-----GlyAlaLysGlnAlaGlyAspValGlySerProG 744  
QY 61 -----CTCAGATCCTGGCGCA 46  
Db ||||| ||| ||||| ||||| :|||  
QY 744 lyAlaMet-AspProGlyArg 750

## RESULT 17

US-08-397-633A-26  
: Sequence 26, Application US/08397633A  
: Patent No. 5773577  
: GENERAL INFORMATION:  
: APPLICANT: Capello, Joseph  
: TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE  
: TITLE OF INVENTION: OF ENZYMATIC CROSS-LINKING  
: NUMBER OF SEQUENCES: 105  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
: STREET: 4 Embarras Center, Suite 3400  
: CITY: San Francisco  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94111-4187  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/397,633A  
: FILING DATE:  
: CLASSIFICATION: 530  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Rowland, Bertram I  
: REGISTRATION NUMBER: 20,015  
: REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (415) 781-1989  
: TELEFAX: (415) 398-3249  
: TELEX: 910 277239  
: INFORMATION FOR SEQ ID NO: 26:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 762 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
US-08-397-633A-26

Alignment Scores:  
Pred. No.: 5,68e-22 Length: 762  
Score: 391.00 Matches: 259  
Percent Similarity: 35.17% Conservative: 41  
Best Local Similarity: 30.36% Mismatches: 296  
Query Match: 6.28% Indels: 259  
DB: 1 Gaps: 51

US-09-759-143-110 (1-3410) x US-08-397-633A-26 (1-762)

QY 2347 GGGAAACCGAGTACTGAGTTTATTTCAGCTCCCAAAACCCCTCTCTAGGTGTCT--- 2291  
Db ||||| ||| ||||| ||||| :|||  
QY 76 GlySerProGly-----AlaProGlyThrProGlyGluGlyGlnGlnHis 90  
QY 2290 CAACTAGGAGGCTAGCTGTTAAACCTGAGCCTGGGTAAATCCACCTGCA-----GAG 2240







QY 782 TGAAGGCATAGACAGAGTAGGCTGGCCACAGTGGT---CCGGGT---CCCGAAGAGGT 729  
||||| ||| ||||| |||||  
Db 534 -----GlyThrProGlnGlyLeuProGlySerProGlyAlaPro 548  
728 CAGAGCAGAGCCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAGTCCAGAGCCCA 669  
||||| :||| ||||| :||| |||||  
Db 549 GlyThrProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPro 566  
668 CGCCAGGATGAGCAGTCCAGGCTCCAGGCGCTGGGATCCGGCAGCAGCCCTCCTA 609  
||| :||| ||||| ||||| |||||  
Db 567 GlyProGlnGlyLeuProGlySerProGlyAla-----ProGlyThrProGlyGluGly 584  
608 GCCAGCGGCGCTTGGGATGAGAAAGA-----GGCTCAGCAGGATGCCCA 564  
||||| ||| ||||| ||||| |||||  
Db 585 GlnGlnHisLeuGlyGlyAlaArgGlnAlaGlySerProGlyAlaPro 604  
583 GlnGlnHisLeuGlyGlyAlaArgGlnAlaGlySerProGlyAlaPro 604  
563 AGGACAGTG---CCAGATGAAGGCGCGCGGCGCCATAGCGTCCACGCCAGTGTGTC 507  
||| :||| ||||| |||||  
Db 605 GlyThrProGlnGlyLeuProGly-----614  
506 TGGCTGAGCTTCTCTCTACCCCTTCCAGCAGCAGAGCGGCGGCACATAGGTGTC 387  
||||| ||| ||||| |||||  
Db 625 -----GlnGlyLeuProGlySerProGlyAla-----633  
386 CTGCGGCGCAACACACCTCCAGCCCAAGGTAGCAGGTGACCCAGCAGAGCTGGGCTT 327  
||| ||||| ||| ||||| |||||  
Db 634 -----ProGlyThrProGlnGlyLeuProGlySerProGlyAlaProGlyThr 651  
326 TCCGGT-----GCCGCGCAGCGCGCTCA 303  
||| ||||| |||||  
Db 652 ProGlyGluGlnGlnHisLeuGlyGlyAlaArgGlnAlaGlySerProGlyAla 671  
302 CCCACAGCTCTGACCATAGTGGCGCAGCGGG---TAGGGTCAAGGGCGGTTCAGG 246  
||| ||| ||| ||||| ||||| |||||  
Db 672 ProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGly 691  
245 CACTCAGAGACTGCTGCTGCTGCTCCAGAGCTGCGCGCTCTCTCTCTCTCTCTG 186  
||||| ||| ||||| ||||| |||||  
Db 691 yThrProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly-----709  
185 CCGCCAACTGCTAGGAATCAGCAGCGCGCCCTTTCTGCCAGCGCTTGGTCCGCGTCC 126  
||| :||| ||||| |||||  
Db 710 -----ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGly 725  
125 AGCTTCTCAGCCCATGCTCAACACCTGCTGCTGGGCGCACCTCAGTGGGCGACAGCTCTC 66  
||| ||| ||||| ||||| ||||| |||||  
Db 725 uGlyGlnGlnHisLeuGly-----GlyAlaArgGlnAlaGlySerProGlyValGly 741  
65 ATCA-----CTCAGATCTCTGGCGCA 46  
||| :||| ||||| |||||  
Db 741 ySerProGlyAlaMet-AspProGlyArg 750

## RESULT 15

US-08-707-237A-84

Sequence 84, Application US/08707237A

Patent No. 5830713

## GENERAL INFORMATION:

APPLICANT: Ferrari, Franco A.

APPLICANT: Capello, Joseph

APPLICANT: Crissman, John W.

APPLICANT: Dorman, Mary A.

TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC

TITLE OF INVENTION: REPETITIVE DNA

NUMBER OF SEQUENCES: 108

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hohbach, Test, Albritton &amp; Herbert

STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California  
COUNTRY: United States  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICANT NUMBER: US/08/707,237A  
FILING DATE: 03-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICANT NUMBER: US 08/175,155  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICANT NUMBER: US 08/053,049  
FILING DATE: 22-APR-1993  
PRIOR APPLICATION DATA:  
APPLICANT NUMBER: US 07/609,716  
FILING DATE: 06-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICANT NUMBER: US 07/269,429  
FILING DATE: 09-NOV-1988  
PRIOR APPLICATION DATA:  
APPLICANT NUMBER: US 07/114,618  
FILING DATE: 29-OCT-1987  
PRIOR APPLICATION DATA:  
APPLICANT NUMBER: US 06/927,258  
FILING DATE: 04-NOV-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-55186-10/WHO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 761 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-707-237A-84

## Alignment Scores:

Pred. No.: 3,99e-22 Length: 761  
Score: 393.00 Matches: 259  
Percent Similarity: 35.17% Conservative: 41  
Best Local Similarity: 30.36% Mismatches: 296  
Query Match: 6.31% Indels: 259  
DB: 2 Gaps: 51

US-09-759-143-110 (1-3410) x US-08-707-237A-84 (1-761)

QY 2347 GGGAAACGAGGAGTGGTATTTATTTCAGCTCCCAAAACCTTCTCTAGTGTGTCT--- 2291  
||| ||||| ||||| ||||| |||||  
Db 75 GlySerProGly-----AlaProGlyThrProGlyGluGlyGlnGlnHis 89  
2290 CAACTAGGAGGCTAGCTGTTAAACCCCTGAGCCCTGAGTAAATCCACCTGCA-----GAG 2240  
||||| ||||| ||||| ||||| |||||  
Db 90 HisLeuGlyGlyAlaGlyGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 109  
2239 TCCCGCGCATTCAGATGCGATGGAGCCCTTCTGGCCCTCCCTGTATAGTCCAGACTGAAC 2180  
||| ||||| ||||| ||||| |||||  
Db 110 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro-----GlyThrProGly 125  
2179 CCCTTGGAGGCTCCAGTCCAGGCGAGCCCTAGAGACTGGGGAGAGAGAGGAGGAGCC 2120  
||| :||| ||||| ||||| |||||  
Db 126 Pro---GlnGlyLeu---ProGlySerPro-----GlyAlaProGlyThrPro 139

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/397,633A  
 FILING DATE:  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Rowland, Bertram I  
 REGISTRATION NUMBER: 20,015  
 REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 781-1989  
 TELEFAX: (415) 398-3249  
 TELEX: 910 277299  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 762 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-397-633A-31

Alignment Scores:	2.8e-22	Length:	762
Pred. No.:	395.00	Matches:	257
Score:	395.00	Conservative:	38
Percent Similarity:	34.50%	Mismatches:	298
Best Local Similarity:	30.06%	Indels:	263
Query Match:	6.35%	Gaps:	52
DB:	1		

US-09-759-143-110 (1-3410) x US-08-397-633A-31 (1-762)

2347	GGGAACCAAGCTGACTGAGTTTATTACGTCCTCCAAACACCCCTCTCTAGGTGTGTCT---	2291
QY		
Db	76 GlySerProGly-----AlaProGlyThrProGlyGluGlnGlnHis 90	
QY	2290 CAACATAGGAGCTAGCTGTTAAACCCCTGAGCTGGGTAATCCACCTGCA-----CAG 2240	
Db		
Db	91 HisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110	
QY	2239 TCCCGCATTCAGTCATGGAGGCTTCTGGCCCTCCCTGTATAGTCCAGACTGAACCC 2180	
Db		
Db	111 GlyProGlnGlyLeuProGlySerPro--GlyAlaPro-----GlyThrProGly 126	
QY	2179 CCCTTGGAAAGCCTCCAGTCAGGACGCCCTAGAGACTGGGACAGAGGAGGACGCC 2120	
Db		
Db	127 Pro--GlnGlyLeu--ProGlySerPro-----GlyAlaProGlyThrPro 140	
QY	2119 CAGCCCCAGCTGTGCACCTACGCACCTCAGCAGCAGCAGGTT-----GGCAGCAGAGAG 2066	
Db		
Db	141 GlyProGlnGlyLeuPro--GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159	
QY	2065 CCACATTACTTTGGCAACACAGAAACTGGCGGCCAGCGCCGACCCCTATGGGGCTAAC 2006	
Db		
Db	160 GlnHisHisLeuGly-----GlyAlaArg 167	
QY	2005 AGGACGGGGAGCTGGGACCCAGTGAGGAGGCCCTCCACCCCAATGTGCTGGAGGTTTT 1946	
Db		
Db	168 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro----- 181	
QY	1945 CTCAGCTGAGTATTTTGGCCCAAGTCGCTCTGTGTCAAATACCTACTGTGTAGCAAAAGTAAAT 1886	
Db		
Db	182 -----GlyProGlnGlyLeuPro----- 187	
QY	1885 GCGCACCAGACCCAGGCTCGGCGACACACCATTATAGCAGTCAGCAGACTGGCTGACCTG 1826	
Db		
Db	188 -----GlySerProGlyAlaProGlyThrProGlyProGln----- 199	
QY	1825 GACAATGAGGCCATAAACAGGATGGGGCCACTGGGACAGCAGAAAGCACTATCCAG 1766	
Db		
Db	200 -----GlyLeuProGlySerProGlyAla-----ProG 209	
QY	1765 GATGCGGAGGTCACGCGAGATGCCCGGC-----CCGGAACCCACTGGGCTC 1718	

QY	1666	GCAGAGCGGGTGGAGGTGGGAGCAGGCCACTGCTCCAGCAGCACCACGTGCCATTAGG	1607
Db	246	oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro-----	261
QY	1606	GAAGGAGCTCCAGGCTTAGGCCTGGCAGGAGCTGGTCATCAGCTGTCTCCTCACTGCT	1547
Db	262	----GlyAlaProGlyThr-----ProGlyProGlnGlyLeu-ProGlySerPro-----	276
QY	1546	AGCACCTCCAGTGTCCCTCGGTATTGGGCAGGAAACACTGCTCTCCCGGTGGTAGAG	1487
Db	277	-----GlyAlaProGlyThrProGlyProGlnG	286
QY	1486	GGAGGCCAGTGTAGGCGCAGGATCTGACGGGCTGAGAAGGTGAACCCGGTAGGCGCGC	1427
Db	286	lyLeuPro-----GlySerProGlyAla-ProGlyThrProGlyGluGly---	300
QY	1426	TGAAGCTGTCCACACGGCCACACTGTGGGACAGGCATGTGGCACCCGCAG-----	1377
Db	301	---GlnGlnHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla	319
QY	1376	-----CCACAGGCAAGACTGCCACACTGGCGCAATAGACTGCTCGAGTGC	1332
Db	320	ProGlyThrProGlyProGlnGly---LeuProGlySerPro-----	332
QY	1331	CGAATCGCTGCACAGCGGCTCCATCACCAGAGAGAACACCGGAGATGGCCACTGCA	1272
Db	333	-----GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro	350
QY	1271	GGAAACA-----GCCCCAGGCTGCCCATCCGAA	1245
Db	351	GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu	370
QY	1244	CGCTTCATCATAGTCTCGGGCCCTCGCTCGCCGCTCAGCTCTGGGCA-----CGC	1191
Db	371	GlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla	390
QY	1190	CTGTGTACAGCCCTCGCCACAGAAATCCGTGTAAACACCGCTGAAGGTCATCAGTGCCA	1131
Db	391	ProGlyThr---ProGlyProGlnGly-----LeuPro	400
QY	1130	TCCAGTGCACACTCAGCCACCAAGACGGCGGCGAGGTGCGGGCATCGCGCAGCAC	1071
Db	401	GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla	420
QY	1070	GCTGTGTACCGCGGGAAGCAGGCGCCCGCCAGGTTC-----GGAAAGCCACAGCGGCCCGC	1014
Db	421	ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly	440
QY	1013	-----ATGGACACCACTGGG	999
Db	441	GluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGly	460
QY	998	GCAGCA--AGGAGGGGCGCAGACGCCCTTCTGCTGGCTCGGTGGGCGCCAGCGCTGCCT	942
Db	461	AlaProGlyThrProGlyProGlnGlyLeu-----ProGlySerProGlyAlaPro	477
QY	941	CTCAGCCACCACTGTGGCTCTACCGCAGGTGAGGAAGATCAGGGTGAGCAGGCCAA	892
Db	478	GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro	497
QY	881	AGAGGCACTCT-----CCTGGGTGCCAGTATAGGGGGCCAGGGCACTGG	837
Db	498	GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis	517
QY	836	TGTCCTCACTCAATGGCAGCAGGA-----GGTAGCCACGACGCCCCCAAGACTGATCA	783
Db	518	LeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro-----	533
QY	782	TGAAGGCATAGACAGTAGTGGCCCTGGCGACAGTGGT---CCGGGT---CCCGAAGAGGT	729
Db	534	-----GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro	548

QY	728	CACAGACAGAGCGCCTCCAGTGGAGTGAAGCACACACCTTGGCCACAGAGTCCAGCAGCCCA	666
Db	549	GlyThrProGlyProGlnGlyLeuProGlySerProGly-----AlaProGlyThrPro	566
QY	668	CGCCAGGATGAGCAGTGCACAGCTCCAGGGCGCTGGGATCCGGGCACACAGCAGCCCTGCTA	609
Db	567	GlyProGlnGlyLeuProGlySerProGlyAla-----ProGlyThrProGlyGluGly	584
QY	608	GCCAGCGCGCCCTCTGGATGAGAAGA-----GGCTCAGCAGAGATGCCCA	564
Db	585	GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro	604
QY	563	AGSACAGTG---CCAGATGAAGCGCGCGGCCCATAGCGTCCACGCCAGTGTGCTAC	507
Db	605	GlyThrProGlyProGlnGlyLeuProGly-----	614
QY	506	TGGCTCAGCCTTAGGAGCGGGACACAGACCGAGCGCCAGCACTGGACCANTGCCAGCACCA	447
Db	615	-----SerProGlyAlaProGlyThrProGlyPro-----	624
QY	446	TGGTCATGAACTTCTCTCTACCCCCACTTCCACGACGACGAGCGGCGCACATAGTGTATCG	387
Db	625	-----GlnGlyLeuProGlySerProGlyAla-----	633
QY	386	CTGGCGCCAAACACACACTCCAGGCGCAAGTTAGCAGGTTGACGACGACAGAGCTGGCGTT	327
Db	634	-----ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr	651
QY	326	TCGGGT-----GCCGACGACGCGGCTCA	303
Db	652	ProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySer	671
QY	302	CCACAGCGCTCTGGACCATAGTGGGCCAGGGCGG---TAGGGCTCAGGGGCGGTTTCAGG	246
Db	672	ProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGly	691
QY	245	CACGCCAGAACTCTCGCTCTCGGCTCTGCTCCAGAAAGCTCGGCGCTCTCTCTCTCTGCTG	186
Db	691	yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly-----	709
QY	185	CGCCCAACTGCCTAGGAATCATCCAGCGCGCCCATTTCTGCCAGCGCTTGGTCCCGGCTC	126
Db	710	-----ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGly	725
QY	125	AGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGACCTCAGTGGGACACAGCTCTC	66
Db	725	uGlyGlnGlnHisHisLeuGly-----GlyAlaArgGlnAlaGlyAspValGly	741
QY	65	ATCA-----CTCAGATCCTGCCCA	46
Db	741	ySerProGlyAlaMet-AspProGlyArg	750
RESULT 14			
US-08-397-633A-31			
; Sequence 31, Application US/08397633A			
; Patent No. 5773577			
; GENERAL INFORMATION:			
; APPLICANT: Cappello, Joseph			
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE			
; TITLE OF INVENTION: OF ENZYMIC CROSS-LINKING			
; NUMBER OF SEQUENCES: 105			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: FLEHR, HONBACH, TEST, ALBRITTON & HERBERT			
; STREET: 4 Embarcadero Center, Suite 3400			
; CITY: San Francisco			
; STATE: California			
; COUNTRY: USA			
; ZIP: 94111-4187			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			

## RESULT 14

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

; APPLICANT: Wang, Aljun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Hepler, William  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
 ; FILE REFERENCE: 210121.427C16  
 ; CURRENT APPLICATION NUMBER: US/09/605,785  
 ; CURRENT FILING DATE: 2000-06-27  
 ; NUMBER OF SEQ ID NOS: 835  
 ; SOFTWARE: FASTSEQ for Windows Version 3.0  
 ; SEQ ID NO 706  
 ; LENGTH: 123  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-605-785-706

Alignment Scores:  
 Pred. No.: 6,16e-25 Length: 123  
 Score: 426.00 Matches: 87  
 Percent Similarity: 97.83% Conservative: 3  
 Best Local Similarity: 94.57% Mismatches: 2  
 Query Match: 6.64% Indels: 0  
 DB: 4 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-605-785-706 (1-123)

QY 1250 ATGGGACGCTGGGCTGTCTCTGCGAGTGGCGCATCTCCCTGGTCTCTCTCTGGTCATG 1309  
 Db 1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20  
 QY 1310 GACCGGCTGGTGCACGCGATTCCGGCACTCGGACAGTCTATTGGCAGTGGCGAGCTTC 1369  
 Db 21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40  
 QY 1370 CCTGTGGTGGCGGTGGCCACATGCGCTGTCCACAGTGTGGCGGTGCACAGTTCAGCC 1429  
 Db 41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAlaSerAla 60  
 QY 1430 GCCCTCACCGGTTCCACCTTCTCAGCCCTGCAGATCTCCCTCACACACTGGCCCTCCTC 1489  
 Db 61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80  
 QY 1490 TACCACCGGAGAGCAGGTGTCTCTGCCCAATAC 1525  
 Db 81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

# RESULT 13

US-08-642-255-120  
 ; Sequence 120, Application US/08642255  
 ; Patent No. 5773249  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CAPPELLO, Joseph  
 ; APPLICANT: FERRARI, Franco A.  
 ; TITLE OF INVENTION: High Molecular Weight Collagen-Like  
 ; TITLE OF INVENTION: Protein Polymers  
 ; NUMBER OF SEQUENCES: 135  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: FLEHR, HOERBACH, TEST, ALBRITTON & HERBERT  
 ; STREET: 4 Embarcadero Center, Suite 3400  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-4187  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/642,255  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:

; NAME: ROWLAND, Bertram I.  
 ; REGISTRATION NUMBER: 20,015  
 ; REFERENCE/DOCKET NUMBER: A55556-3/BIR  
 ; TELEPHONE: (415) 494-8700  
 ; TELEFAX: (415) 494-8771  
 ; TELEX: 910 277299 FHT UR  
 ; INFORMATION FOR SEQ ID NO: 120:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 762 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-642-255-120

Alignment Scores:  
 Pred. No.: 2.8e-22 Length: 762  
 Score: 395.00 Matches: 257  
 Percent Similarity: 34.50% Conservative: 38  
 Best Local Similarity: 30.06% Mismatches: 298  
 Query Match: 6.35% Indels: 263  
 DB: 1 Gaps: 52

US-09-759-143-110 (1-3410) x US-08-642-255-120 (1-762)

QY 2347 GGGAAACAGGTGACTGAGTTTATTTCAGCTCCCAAAACCCCTCTCTAGTGTGTCT--- 2291  
 Db 76 GlySerProGly-----AlaProGlyThrProGlyGluGlnGlnHis 90  
 QY 2290 CACTAGGAGGTAGCTGTTAACCCCTGAGCTGGGTAAATCCACCTGCA-----GAG 2240  
 Db 91 HisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110  
 QY 2239 TCCCGCATTCAGTGCATGGAGCCCTCTCGGCTCCCTGTATAGTCCAGACTGAACCC 2180  
 Db 111 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro-----GlyThrProGly 126  
 QY 2179 CCCTTGAAGGCTCCAGTCAGCAGCCCTAGAGACTGGGGAGAGAGAGAGAGAGAGAGAG 2120  
 Db 127 Pro---GlnGlyLeu---ProGlySerPro-----GlyAlaProGlyThrPro 140  
 QY 2119 CAGCCCGGAGCTGTGAGCTACCCACTCAGCAGCAGAGGT-----GGCAGCAGAGAG 2066  
 Db 141 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGln 159  
 QY 2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGCCAGCCCGCAGCCCGCCAGGCTAAC 2006  
 Db 160 GlnHisHisLeuGly-----GlyAlaArg 167  
 QY 2005 AGGAGCGGGAGCTGGGAGCCAGTGAGGAGCGCCCTCCACCCCAATGTCTGGAAGTTT 1946  
 Db 168 GlnAlaGlyAspValGlySerPro---GlyAlaProGlyThrPro-----181  
 QY 1945 CTACGCTGAGTATTGGCCAAGTCGCTCTGTGCAATATACCTGTGTAGCAAGTAAT 1886  
 Db 182 -----GlyProGlnGlyLeuPro-----187  
 QY 1885 GCGGACCGAGCCAGCCCTGGCGGAGACACCATATAGGCAAGTACAGACTGCTGAGCTG 1826  
 Db 188 ---GlySerProGlyAlaProGlyThrProGlyProGln-----199  
 QY 1825 GACANTGGAGCCCATAAACAGGGATGGGGCCACCTGGGACACGACGAGGAGGCACTTCAG 1766  
 Db 200 -----GlyLeuProGlySerProGlyAla-----ProG 209  
 QY 1765 GATGGCGAGGTCCAGGCAGATGCCCGGC-----CCGAAACCCAGCTGGCCTC 1718  
 Db 209 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 228  
 QY 1717 GGTGGGCTCACCACACACACACGTCACGGAGACATCACAGGCA-----GAGGCCCC 1667  
 Db 229 Gly-----GlnGlnHisHisLeuGlyGly-AlaArgGlnAlaGlyAspValGlySerPr 246

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6252047e  
US-09-525-397-36

Alignment Scores:  
Pred. No.: 7e-91  
Score: 1287.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 20.05%  
Indels: 0  
Gaps: 4

US-09-759-143-110 (1-3410) x US-09-525-397-36 (1-255)

```
Qy 1178 GGCGTGTACAGCGGTGCCCCAGAGTGGAGCGGGGACCGAGCGCCGGGAGACACTATGAT 1237
Db 1 GlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAsp 20
Qy 1238 GAAGGCGTTCGGATGGCGACGCTGGGCTGTTCTTCAGTGGCGCCATCTCCCTGGTCTTC 1297
Db 21 GluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPhe 40
Qy 1298 TCTGTGTCATGACCGGCTGGTGCACCGATTCCGCACTCGAGCAGTCTATTGGCCAGT 1357
Db 41 SerLeuValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSer 60
Qy 1358 GTGGCAGCTTTCCTGTGGCTGCGGCTGCGGTCACATGCTGTCACAGTGGCGCGGNG 1417
Db 61 ValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaVal 80
Qy 1418 ACAGCTTCAGCGCGCTTCACCGGTTTCACCTTCTCAGCCCTGCAGATCTCCCTCACAC 1477
Db 81 ThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnLeuProTyrThr 100
Qy 1478 CTGCGCTCCTCTACACCGGAGACAGAGTGTTCCTGCCAAATACCGAGGGACACT 1537
Db 101 LeuAlaSerLeuTyrHisArgGluGlnValPheLeuProLysTyrArgGlyAspThr 120
Qy 1538 GGAGTGCTAGCAGTGGAGGACAGCTGATCAGCAGCTTCCTGCGAGCCCTTAAGCCTGGA 1597
Db 121 GlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGly 140
Qy 1598 GTCCTCTCCCTAATGACACGCTGGTCTGAGGACAGTGGCTGCTCCACCTCCACCC 1657
Db 141 AlaProPheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProPro 160
Qy 1658 CGCTCTGCGGGGCTCTGCTGTGATGTCTCCGTAGTGTGGTGGGTGGAGCCACC 1717
Db 161 AlaLeuCysGlyAlaSerAlaCysAspValSerValArgValValGlyGluProThr 180
Qy 1718 GAGGCCAGGGTGTTCGGGGCGGGGATCTGCTGCGACCTCGCCATCTCGATAGTGC 1777
Db 181 GluAlaArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAla 200
Qy 1778 TTCCTGTGTCCTCCAGGTGGCCCCATCCCTGTTTATGGCTCCATGTTCAGCTCAGCCAG 1837
Db 201 PheLeuLeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGln 220
Qy 1838 TCTGTACTGCTATATGGTGTCTGCGCGAGGCTGGTGTGGTGGCTTACTTGTCT 1897
Db 221 SerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAla 240
Qy 1898 ACAGAGTACTATTGACAGAGGACTTGCCCAATACTCAGCG 1942
Db 241 ThrGlnValValPheAspLysSerAspLeuAlaLysTyrSerAla 255
```

## RESULT 11

US-09-439-313-571  
; Sequence 571, Application US/09439313;  
; Patent No. 6329505  
; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang Yuqi  
APPLICANT: Reed, Steven G.  
APPLICANT: Kalos, Michael  
APPLICANT: Fanger, Gary  
APPLICANT: Retter, Mark  
APPLICANT: Solk, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
FILE REFERENCE: 210121.427C9  
CURRENT APPLICATION NUMBER: US/09/439,313  
CURRENT FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 571  
LENGTH: 84  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-439-313-571

Alignment Scores:  
Pred. No.: 5.45e-27  
Score: 452.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 7.04%  
Indels: 0  
Gaps: 4

US-09-759-143-110 (1-3410) x US-09-439-313-571 (1-84)

```
Qy 1493 CACCGGAGAGCAGGTGTCTGCTGCCAAATACCGAGGGACACTGGAGGTCTAGCAGT 1552
Db 1 HisArgGluGlnValPheLeuProLysTyrArgGlyAspThrGlyAlaSerSer 20
Qy 1553 GAGGACAGCTGATGACACAGCTTCTGCGAGCCCTAACCTGGAGCTCCCTCCCTAAT 1612
Db 21 GluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPheProAsn 40
Qy 1613 GGACAGTGGTGGAGGAGTGGCTGCTCCACCTCCACCTCCCGGCTCTCGCGGGCC 1672
Db 41 GlyHisValGlyAlaGlySerGlyLeuLeuProProProAlaLeuCysGlyAla 60
Qy 1673 TCTGCTGTGATGTCTCCGTACGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1732
Db 61 SerAlaCysAspValSerValArgValValGlyGluProThrGluAlaArgValVal 80
Qy 1733 CCGGGCGGGG 1744
Db 81 ProGlyArgGly 84
```

## RESULT 12

US-09-605-785-706  
; Sequence 706, Application US/09605785  
; Patent No. 6321716  
; GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolck, John H.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrik  
APPLICANT: Li, Samuel





Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
Qy 1844 ACTGCCTATATGCTCTCCGAGCCCTGGGTGGTGGTCTTACTTCTGCTACAG 1903  
Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
Qy 1904 GTAGTATTTCACAGAGCAGCTTGGCCAAATACTACAGC 1942  
Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

## RESULT 8

US-09-605-785-708

: Sequence 708, Application US/09605785

: Patent No. 6321716

: GENERAL INFORMATION:

: APPLICANT: Xu, Jiangchun

: APPLICANT: Dillon, Davin C.

: APPLICANT: Mitcham, Jennifer L.

: APPLICANT: Harlocker, Susan L.

: APPLICANT: Jiang, Yuqi

: APPLICANT: Henderson, Robert A.

: APPLICANT: Kalos, Michael D.

: APPLICANT: Fanger, Gary R.

: APPLICANT: Retter, Marc W.

: APPLICANT: Stolk, John A.

: APPLICANT: Day, Craig H.

: APPLICANT: Vedvick, Thomas S.

: APPLICANT: Carter, Darrick

: APPLICANT: Li, Samuel

: APPLICANT: Wang, AiJun

: APPLICANT: Skeiky, Yasir A.W.

: APPLICANT: Hepler, William

: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

: TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

: FILE REFERENCE: 210121.427C16

: CURRENT APPLICATION NUMBER: US/09/605,785

: CURRENT FILING DATE: 2000-06-27

: NUMBER OF SEQ ID NOS: 835

: SOFTWARE: FastSeq for Windows Version 3.0

: SEQ ID NO 708

: LENGTH: 371

: TYPE: PRT

: ORGANISM: Homo sapiens

US-09-605-785-708

## Alignment Scores:

Pred. No.:	4.58e-101	Length:	371
Score:	1420.50	Matches:	275
Percent Similarity:	88.58%	Conservative:	12
Best Local Similarity:	84.88%	Mismatches:	23
Query Match:	22.13%	Indels:	14
DB:	4	Caps:	1

US-09-759-143-110 (1-3410) x US-09-605-785-708 (1-371)

Qy. 347 AACCTGTAACCTTTGGCCTGGAGGTGTGTTGGCCGAGGATACACCTATGTGCCGCT 406  
Db 9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28  
Qy 407 CTGCTGCTGGAGTGGGGGTAGAGGAGAGTTATGACCATGGTGGCGCATTTGTCCA 466  
Db 29 LeuLeu-----AlaGlyIleGlyPro 35  
Qy 467 GTGCTGGCGCTGTCTGTCTCCGCTCTTAGCTACAGCAGTACACCTGGCGGTGGACGC 526  
Db 36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55  
Qy 527 TATGCGCGCGCGCGCGCTTCACTGCGGCACTGTCTTGGGCACTCTGCTGGCGCTCTTT 586  
Db 56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75  
Qy 587 CTCATCCAGGCGCGCTGCTAGCAGGCTGCTGTGCCCGGATCCAGCGCCCTGGAG 646

Db 76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95  
Qy 647 CTGGCACTGCTCATCTGGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706  
Db 96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115  
Qy 707 CCACCTGGAGGCGCTCTCTGACCTCTTCCGGGACCCGCGACCTCTGCGCAGCGCTAC 766  
Db 116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135  
Qy 767 TCTGTCTATGCTTCATCATCTGCTGGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCT 826  
Db 136 SerValTyrAlaPheMetIleSerLeuGlyCysLeuGlyTyrLeuLeuProAlaIle 155  
Qy 827 GACTGGGACACAGTGGCCCTGCGCCCTACCTGGGACCCAGGAGGAGTGTCTTTGGC 886  
Db 156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPheGly 175  
Qy 887 CTGCTCACCTCATCTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 946  
Db 176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195  
Qy 947 GCGCTGGCGCCACCGACCGGCTGTGGGCGCCCTCTCTGCTGCGCCCTGCTGCTGCT 1006  
Db 196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215  
Qy 1007 TGTCCATGCGGGCGCGCTTGGCTTTCGGGAACCTGGGCGCGCTGCTTCCCGGCTGCAC 1066  
Db 216 CysProCysArgAlaAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235  
Qy 1067 CAGCTGTGCTGCGCATGCGCGCACCTGCGCGCGCTGCTTCTGCTGCTGCTGCTGCTGCT 1126  
Db 236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255  
Qy 1127 TGGATGGCACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1186  
Db 256 TrpMetAlaLeuMetPheThrLeuPheTyrThrAspPheValGlyGlyLeuTyr 275  
Qy 1187 CAGGGCGTCCCGAGCTGAGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGGCTT 1246  
Db 276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295  
Qy 1247 CGATGGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1306  
Db 296 AlaLeuAlaAlaSerArgGlyTrpCysGlySerArgProGluThrThrLeuGlyAla 315  
Qy 1307 ATGGACCGGC 1316  
Db 315 aValSerGly 318

## RESULT 9

US-09-071-710-36

: Sequence 36, Application US/09071710

: Patent No. 6130043

: GENERAL INFORMATION:

: APPLICANT: BILLING-MEDEL, PATRICIA

: APPLICANT: COHEN, MAURICE

: APPLICANT: COLPITTS, TRACEY L.

: APPLICANT: FRIEDMAN, PAULA N.

: APPLICANT: GORDON, JULIAN

: APPLICANT: GRANADOS, EDWARD N.

: APPLICANT: HODGES, STEVEN C.

: APPLICANT: KLASS, MICHAEL R.

: APPLICANT: KRATOCHVIL, JON D.

: APPLICANT: ROBERTS-RAPP, LISA

: APPLICANT: RUSSELL, JOHN C.

: APPLICANT: STROUPE, STEPHEN D.

: TITLE OF INVENTION: REAGENTS AND METHODS USEFUL

: TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE

: NUMBER OF SEQUENCES: 41

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Abbott Laboratories



QY 1724 AGGTGGTCCGGCGGCGCATCTGCTGGACCTCCCATCTGATAGTGCCTTCTG 1783  
|||||  
Db 481 ArgValValProGlyArgGlyIleCysLeuAspAlaIleLeuAspSerAlaPheLeu 500  
|||||  
QY 1784 CTGTCCAGGTGGCGGCGCATCTGTTATATGGGCTCCATCTCCAGCTCAGCCAGTGTG 1843  
|||||  
Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
|||||  
QY 1844 ACTGCCATATATGTGTGTCGCCGAGGCGTGGTCTGCTGCCCATTTACTTTGCTACACAG 1903  
|||||  
Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleIleIlePheAlaThrGln 540  
|||||  
QY 1904 GTAGTATTTGACAGAGCGGACTTGGCCAAATACTCAGCG 1942  
|||||  
Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553  
|||||

## RESULT 6

US-09-602-877A-101  
; Sequence 101, Application US/09602877A  
; Patent No. 6432707  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.446C5  
; CURRENT APPLICATION NUMBER: US/09/602,877A  
; CURRENT FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 101  
; LENGTH: 553  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-602-877A-101

Alignment Scores:  
Pred. No.: 1,61e-211 Length: 553  
Score: 2861.00 Matches: 553  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 44.58% Indels: 0  
DB: 4 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-602-877A-101 (1-553)

QY 284 ATGTCTCAGAGGTGTGGGTGAGCGGCTGCTGGCGACCGGAAAGCCAGCTTCTGCTG 343  
|||||  
Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
|||||  
QY 344 GTCAACCTGCTAACCTTTGGGCTGTGGAGTGTGTTGGCGGCGAGGCATCATTATGTGCCG 403  
|||||  
Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyrValPro 40  
|||||  
QY 404 CTTCTGCTGCTGGAAGTGGGGGTAGAGGAGAGTTGACCATGCTGGCTGGGCAATTGGT 463  
|||||  
Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
|||||  
QY 464 CCAGTGTGGGCTGTGCTGCTCCGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 523  
|||||  
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
|||||  
QY 524 CCTATGGCGCGCGCGGCTTTCATCTGGGCACTGTCTTGGGCACTCTCTGCTAGGCTC 583  
|||||  
Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
|||||  
QY 584 TTCTCATCTCCAAAGCGCGGTGTGTAGAGGCTGTGTGTGCGCGGATCCAGGCGCCCTG 643  
|||||  
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
|||||  
QY 644 GAGTGGGCACTGCTCATCTCTGGGCGTGGGCTGCTGGACTTCTGTGGCGCAGGTGTGCTTC 703  
|||||

Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
|||||  
QY 704 ACTCCACTGGAGGCGCTCTCTACCTCTTCCGGGACCGGACACTGTCTCGCAGGCC 763  
|||||  
Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
|||||  
QY 764 TACTCTGTCTATGCTTTCATGATCAGTCTTGGGGCTGCTGGGCTACCTCTCTGCTGCC 823  
|||||  
Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
|||||  
QY 824 ATTGACTGGACACAGTGCCTTGGCCCTTACCTTGGGACCCAGGAGGAGTGCCTCTTT 883  
|||||  
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200  
|||||  
QY 884 GGCCTGCTCACCTCATCTTCTCCTACCTGCTGAGGACACACTGCTGCTGCTGCTGAGG 943  
|||||  
Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
|||||  
QY 944 GCAGGCTGGGCGCCACCGAGCCAGAGAGGGCTGTGGGCGCCCTCTCTTGTGCGCCGAC 1003  
|||||  
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
|||||  
QY 1004 TGCTGCTCATGCGGCGGCGCTTGGCTTCCGGAACCTGGGCGCCCTGCTTCCCGGCTG 1063  
|||||  
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
|||||  
QY 1064 CACCAGCTGTGCTGGCGCATGCGCGCACCTGCGCGGCTCTTCTGCTGCTGCTGCTG 1123  
|||||  
Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgPheValAlaGluLeuCys 280  
|||||  
QY 1124 AGCTGATGGCACTCATGACCTTCACGCTGCTTTACACGGATTTCTGCTGGCGGCGCTG 1183  
|||||  
Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
|||||  
QY 1184 TACGAGGCGTGGCGGAGCTGAGCGGCGCACCGAGCGCGGAGAGACACTATGATGAAGC 1243  
|||||  
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgHisTyrAspGluGly 320  
|||||  
QY 1244 GTTTCGGATGGGAGCTGGGCTGTCTGCTGAGTGGCGCATCTCTCTGCTCTCTCTG 1303  
|||||  
Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
|||||  
QY 1304 GTCATGGACCGGCTGCTGGCGGATTCGGCACTCGAGCACTGCTATTTGGCGGAGTGGCA 1363  
|||||  
Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
|||||  
QY 1364 GCTTTCCTGCTGGCTGCGGTCACATGCCTGTCTCCACAGTGTGGCGCTGTGACAGCT 1423  
|||||  
Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380  
|||||  
QY 1424 TCAGCGCGCTCACCGGCTTCACCTTCTCAGCCCTCAGATCTGCTGCTACACACTGGCC 1483  
|||||  
Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnLeuProTyrThrLeuAla 400  
|||||  
QY 1484 TCCTCTTACCACCGGAGAGAGGTGTCTCTCCCAAAATACCGAGGAGGACACTGGAGGT 1543  
|||||  
Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
|||||  
QY 1544 GCTAGCAGTGGAGACACCTGATGACACAGCTTCTCTGCGAGGCGCTTAAAGCTGAGCTCC 1603  
|||||  
Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
|||||  
QY 1604 TTCCCTAATGGACACGTGGGTGTGGAGGAGGTGGCTGCTGCCACTCCACCGCGGCTC 1663  
|||||  
Db 441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 460  
|||||  
QY 1664 TGGCGGCGCTCTGCTGATGTCTCTCTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1723  
|||||  
Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
|||||  
QY 1724 AGGTGTGTTCGGGCGCGGCGATCTGCTGAGCTGCTGCCATCTGCTGCTGCTGCTGCTGCT 1783  
|||||







Db 541 valValPheAspLysSerAspLeuAlaLysTyrSerAla 553

## RESULT 3

US-09-785-113  
Sequence 113, Application US/09605785  
Patent No. 6321716

## GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C16  
CURRENT APPLICATION NUMBER: US/09/605,785  
CURRENT FILING DATE: 2000-06-27  
NUMBER OF SEQ ID NOS: 835  
SOFTWARE: FastSeq for Windows version 3.0  
SEQ ID NO 113  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-605-785-113

Alignment Scores:  
Pred. No.: 1,616-211 Length: 553  
Score: 2861.00 Matches: 553  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 44.58% Indels: 0  
DB: 4 Gaps: 0  
US-09-759-143-110 (1-3410) x US-09-605-785-113 (1-553)

QY 284 ATGGTCACAGAGCTGTGGGTGAGCGCTGTGGGACCGAAGCCAGCTCTTGCTG 343  
Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
QY 344 GTCAACCTGTAACTTTGGCTGGAGGTGTGGCGCGCAGCATCACCTATGTCGG 403  
Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyrValPro 40  
QY 404 CCTCTGCTGTGGAAGTGGGGGTAGAGGAGAAGTTATGACCATGGTGTGGCATTTGGT 463  
Db 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
QY 464 CCAGTGTGGCGCTGGTCTGTCTCCGCTCTAGCTCAGCCATGACCATGTCGGCGTGA 523  
Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
QY 524 GCCTATGGCGCGCGCGCTTCTATCTGGGCACTGTCTTGGGCATCTTGGCTGAGGCTC 583  
Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
QY 584 TTTCTCATCCAAAGGGCGGTGTAGCAGGGCTGTGTGCCCGGATCCCGAGCCCTTG 643  
Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
QY 644 GAGCTGGCACTGCTCATCTGGGCGTGGGCTGTCTGGACTTCTGTGGCCAGGTGTCTC 703  
Db 1

Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
QY 704 ACTCCACTGGAGCGCTCTCTGACCTCTTCGGGACCGACCGACCACTGTGGCAGGCC 763  
Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
QY 764 TACTCTGTCTATGCTTCATCATCATCTTGGGGGCTGGCTGGCTGCTCTCTCTCTCT 823  
Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyCysLeuGlyTyrLeuLeuProAla 180  
QY 824 ATTGACTGGGACACCACTGTCCTGGCCCTTACCTGGGCACCCAGGAGGAGTGTCTCT 883  
Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200  
QY 884 GGCCTGTCTACCTTCCTCTCTGCTGTAGCAGCCACACACACACACACACACACAC 943  
Db 201 GlyLeuLeuThrIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
QY 944 GCAGCGCTGGGCGCCACCGAGCCAGCAGAGGGCTGTGGCCCTCTCTCTCTCTCTCT 1003  
Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
QY 1004 TGCTGTCCATGCGGGCGCGCTTGGCTTCCGGAACCTGGGCGCCCTCTCTCTCTCTCT 1063  
Db 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
QY 1064 CACCAGCTGTGTGGCGCATGCCCGCACCTGGCGCGCTCTCTCTCTCTCTCTCTCT 1123  
Db 261 HisGlnLeuCysArgMetProArgThrLeuArgLeuPheValAlaGluLeuLeuCys 280  
QY 1124 AGCTGTGATGCACTCATGACCTTACAGCTGTTCACAGGATTTCTGGCGGAGGGCTG 1183  
Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
QY 1184 TACCAGGGCGTCCCGAGAGCTGAGCGCGGACCGAGCGCGGAGACACTATGATGAAGGC 1243  
Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
QY 1244 GTTCGGATGGGACGCTGGGCTGTTCCTGACGTGGCCCATCTCCCTGGTCTCTCTCTCT 1303  
Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
QY 1304 GTCATGGACCGCTGGTGTGCGAGCATTCGCGACACTCGAGCAGTCTATTGGCCAGTGTG 1363  
Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
QY 1364 GCTTTTCCCTGTGGCTGCGGTGCGACATGCTCTCCACAGTGTGGCGCTGGTGTGACAGCT 1423  
Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380  
QY 1424 TCAGCGCGCTCACCGGGTTACCTTCTCAGCCCTGCAGATCTGCTCTACACACTGGCC 1483  
Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
QY 1484 TCCTCTACACCGGAGAGAGGTTCTCTGCGCCCAATACCGAGGGGACACTGGAGGT 1543  
Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
QY 1544 GCTAGCAGTCAGGACAGCCTGTGACACAGCTTCTGCGAGCGCTTAAGCCTGAGGCTCC 1603  
Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
QY 1604 TTCCCTTAATGGACACCTGGGTGCTGGAGCAGTGGCCTGCTCCCACTCCACCCCGGCTC 1663  
Db 441 PheProAsnGlyHisValGlyAlaGlySerGlyLeuLeuProProProProAlaLeu 460  
QY 1664 TGGGGGCGCTCTGCTGTGATCTCCCTACCTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1723  
Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
QY 1724 AGGGTGTTCGGGGCGGGGCGATCTGCTGTGGACCTGCGCATCTCTGGATGAGTGTCTCTG 1783  
Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500



MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatenCin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/030,607  
 FILING DATE: 25-FEB-1998  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Makl, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.427C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 113:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 553 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 US-09-030-607-113

Alignment Scores:  
 Pred. No.: 1,61e-211 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 4 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-030-607-113 (1-553)

QY 284 ATGTTCAGAGGCTGGGTGAGCCGGCTGCTGGCGGACCGGAAAGCCAGCTCTGCTG 343  
 DB 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
 QY 344 GTCACCTGCTAACCTTTGGCGCTGGAGGTGTGTTGGCGGACGAGCATACCTATGTCCG 403  
 DB 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40  
 QY 404 CCTCTGCTGCTGAAGTGGGGGTAGAGGAGAGTTCATGACCAATGGTGTGGGCAATGGT 463  
 DB 41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60  
 QY 464 CCAGTGTGGGCTGCTGTGTGTCGGCTCTAGGCTCAGCCAGTGCACCTGGCGTGA 523  
 DB 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CGCTATGGCGCGCGCCCTTCATCTGGGCACTGCTTGGGCATCCTGCTGAGCCTC 583  
 DB 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCCAAGCGCGCTGGCTAGCAGGCTGTGTGCCGGATCCAGGCCCTG 643  
 DB 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120  
 QY 644 GAGCTGGACATGCTCATCTGCTGGGCTGGGCTGCTGACTTCTGTGGCCAGGTGCTTC 703  
 DB 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCCACTGGAGGCGCTCTCTGACCTCTTCGGGACCGCGGACCACTGTGCGCAGGCC 763  
 DB 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160  
 QY 764 TACTCTGTATGCTTCATCATCATGCTTGGGGGCTGCTGGGCTACCTCTGCTGCTGCC 823  
 DB 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATGACTGGGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 883

DB 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200  
 QY 884 GGCCTGCTCACCTCATCTTCTCAGCTGGGTAGCAGCCACACCTGCTGGTGGCTGAGGAG 943  
 DB 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 GCAGCGCTGGCGCCACCGAGCAGCAGAGAGGGCTGTGCGGCCCTCTTGTGCGCCCGAC 1003  
 DB 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGCTGTCCATGCGCGCGCGCTGGCTTTCGGAACTTGGGGGCGCTGCTTCCCGGGCTG 1063  
 DB 241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACCAGCTGTGCTGCGCATGCCCGCACCTGCGCGCTTCTGCTGCTGCTGCTGCTGCTG 1123  
 DB 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACAGGATTTCTGGGCGAGGGCTG 1183  
 DB 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
 QY 1184 TACCAGGCGTCCCGAGAGCTGAGCGGCGCAGGCGCGGAGGCGGAGACACTATGATGAAGC 1243  
 DB 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320  
 QY 1244 GTTCGATGGCAGCCTGGGCTGTTCCTGCACTGCGGCACTCTGCTGCTGCTGCTGCTGCTG 1303  
 DB 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTCATGGACCGGCTGGTGTGAGGATTCGCACTCGAGCAGTCTATTTGGCCAGTGGCA 1363  
 DB 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 QY 1364 GCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423  
 DB 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla 380  
 QY 1424 TCAGCGCGCTCACCGGGTTCACCTTCTCAGCCTCGACATCTGCGCTTACACACTGGCC 1483  
 DB 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnLeuProTyrThrLeuAla 400  
 QY 1484 TCCTCTACCGCGGAGCAGCTGTGACAGCTTCTGCGGCAATACCGAGGCGACACTGGAGGT 1543  
 DB 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
 QY 1544 GCTAGCAGTGAGGACAGCTGTGACAGCTTCTGCGGCGCTTAAAGCTTGAAGCTGCC 1603  
 DB 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 QY 1604 TCCCTTAATGAGCAGCTGGTGTGAGGAGTGGCTGCTGCCACTTAAAGCTTGAAGCTGCC 1663  
 DB 441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 460  
 QY 1664 TGGCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1723  
 DB 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
 QY 1724 AGGCTGTGCTGGCGCGGCGCATCTGCTGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCTG 1783  
 DB 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGTCCAGGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1843  
 DB 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTGCTCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1903  
 DB 521 ThrAlaTyrMetValSerAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTGACAGGAGGAGCTTGGCCAAATACTCAGC 1942

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-09-020-956-113

## Alignment Scores:

Pred. No.: 1,61e-211 Length: 553  
Score: 2861.00 Matches: 553  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 44.58% Indels: 0  
DB: 4 Gaps: 0

US-09-759-143-110 (1-3410) x US-09-020-956-113 (1-553)

QY	284	ATGGTCCAGAGCTGGGTGAGCGCCCTGCTGGCGCACCAGGAAAGCCAGCTCTGCTG	343
DB	1	MetValGlnArgLeuTrpValSerArgLeuLeuAArgHisArgLysAlaGlnLeuLeuLeu	20
QY	344	GTCACACTGCTAACCTTTGGCTGGAGGTGCTTTGGCGCAGGACATCACCTATGTGCGC	403
DB	21	ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro	40
QY	404	CCTCTGCTGCTGAAGTGGGGGTAGAGGAGAGTTTCATGACCATGGTCTGGGCAATGGT	463
DB	41	ProLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly	60
QY	464	CCAGTCTGGGCTGTGCTGTCTCCCGCTCCCTAGGCTAGGCTAGGCTAGGCTGGCGTGA	523
DB	61	ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly	80
QY	524	CGCTATGGCGCGCGCGCTTCATCTGGGCACTGTCTTGGGCACTCTGCTGGAGCTC	583
DB	81	ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu	100
QY	584	TTTCTCATCCCAAGGCGCGGTGGCTAGCAGGCTGTCTGCGCGAGTCCAGGCCCTCG	643
DB	101	PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu	120
QY	644	GAGCTGGCACTGTCTATCTGGCGGTGGGCTGTCTGGCTCTGTGGCGAGGTGCTTC	703
DB	121	GluLeuAlaLeuLeuLeuLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe	140
QY	704	ACTCCACTGGAGCCCTGCTCTCTGACCTCTTCCGGGACCCGACCACTGTCCGACGCC	763
DB	141	ThrProLeuGluAlaLeuLeuSerAspPheArgAspProAspHisCysArgGlnAla	160
QY	764	TACTCTGTATGCTTTCATGATCAGTCTTGGGGCTGTCTGGGCTACCTCTCTGCTGCC	823
DB	161	TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla	180
QY	824	ATTGACTGGACACCACTGCTGGCGCTTCTACCTGGGCACTGGGCGCTGCTCCCTCTT	883
DB	181	IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe	200
QY	884	GGCGTGTCTACCTCATCTTCTCTACCTGCTAGCAGCCACACTGCTGTGCTGAGGAG	943
DB	201	GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu	220
QY	944	GCAGCGTGGCGCCCGCCAGGAGGCTGTGGCGCCCTCTGCTGCTGCTGCTGCTGCTG	1003
DB	221	AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis	240
QY	1004	TGCTGTCCATGCGCGCGCTTGGCTTCCGGAACCTGGGCGCTGCTTCCCGGCTG	1063
DB	241	CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu	260
QY	1064	CACAGCTGTGCTGCGCGATGCCCCGACCCCTGCGCGGCTCTCTGCTGCTGCTGCTG	1123
DB	261	HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys	280

QY	1124	AGCTGGATGCACCTCATGACCTTCACGCTCTTTTACACGGATTTCGTGGCGGAGGCGTG	1183
DB	281	SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu	300
QY	1184	TACCAGGGCTGCCAGAGCTGAGCGGGCAGCGAGGCGCGGAGACACTATGATGAAGCC	1243
DB	301	TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgHisTyrAspGluGly	320
QY	1244	GTTCCGATGGCAGCCCTGGGGCTTCTCTGACGTGGGCCCATCTCCCTGGTCTTCTCTCTG	1303
DB	321	ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu	340
QY	1304	GTCATGACCGGTGGTGGCAGGATTCGGCAGCTCGGACAGCTCTATTGGCCCATGTGGCA	1363
DB	341	ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla	360
QY	1364	GCTTTCCCTGCTGGCTGCCGTGGCCACATGCTGTCCACAGTGTGGCGGTGACAGCT	1423
DB	361	AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla	380
QY	1424	TCAGCCGCCCTCACCGGGTTTCACCTTCTCAGCCCTGCAGATCTCCCTACACACTGGCC	1483
DB	381	SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla	400
QY	1484	TCCCTCTACACCGCGGAGAGCAGGTGCTTCTGCCCAANTACCGAGGGGACACTGGAGT	1543
DB	401	SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly	420
QY	1544	GCTAGCAGTGGAGGACGCTGATGACGAGCTTCTGCGCAGGCTTAAAGCTGGAGCTCCC	1603
DB	421	AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro	440
QY	1604	TTCCCTTAATGGACAGCTGGGTGGTGGAGGAGTGGCTTCCACCTCCACCTCCAGCGCTC	1663
DB	441	PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu	460
QY	1664	TGCGGGGCTCTGCTGTGATGCTCCGTAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1723
DB	461	CysGlyAlaSerAlaCysAspValSerValArgValValGlyGluProThrGluAla	480
QY	1724	AGGTGGTTCGCGGCGCGGCTGCTGCTGACCTGCGACCTCCCTCTGGATAGTGGCTTCTG	1783
DB	481	ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu	500
QY	1784	CTGTCAGGAGTGGCGCCATCCCTCTTATGGGCTCCATTTGTCAGCTACGCTGCTGTC	1843
DB	501	LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal	520
QY	1844	ACTGCTATATGGTGTCTGCGCAGGCTGGGTGCTGGTGGCTATTTACTTTGTACACAG	1903
DB	521	ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln	540
QY	1904	GTAGTATTGACAGAGCGACTTGGCCAAATPACTACGCG	1942
DB	541	ValValPheAspLysSerAspLeuAlaLysTyrSerAla	553

## RESULT 2

US-09-030-607-113  
; Sequence 113, Application US/09030607  
; Patent No. 6262245

## GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS  
NUMBER OF SEQUENCES: 224  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 9, 2003, 22:23:09 ; Search time 29.5 Seconds  
(without alignments)  
6802.181 Million cell updates/sec

Title: US-09-759-143-110  
Perfect score: 6418  
Sequence: 1 ggggaaccagctgcacgcgc.....aaaaataaaaaaaaaa 3410

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_n2p.model  
-Q=/cgn2\_1/USPTO\_spool/US09759143/runat\_05062003\_111320\_23081/app\_query.fasta\_1.3591  
-DB=Issued Patents AA -QFWT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=45  
-MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09759143@cgn\_1\_1\_26#runat\_05062003\_111320\_23081 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	2861	44.6	553 4	US-09-020-956-113
2	2861	44.6	553 4	US-09-030-607-113
3	2861	44.6	553 4	US-09-605-785-113
4	2861	44.6	553 4	US-09-439-313-113
5	2861	44.6	553 4	US-09-352-616A-113
6	2861	44.6	553 4	US-09-602-877A-101
7	2861	44.6	553 4	US-09-232-149A-113
8	1420.5	22.1	371 4	US-09-605-785-708
9	1287	20.1	255 4	US-09-071-710-36
10	1287	20.1	255 4	US-09-525-397-36
11	452	7.0	84 4	US-09-439-313-571
12	426	6.6	123 4	US-09-605-785-706

c 13	395	6.3	762	1	US-08-642-255-120	Sequence 120, App
c 14	395	6.3	762	1	US-08-397-633A-31	Sequence 31, Appl
c 15	393	6.3	761	2	US-08-707-237A-84	Sequence 84, Appl
c 16	391	6.3	762	1	US-08-642-255-114	Sequence 114, App
c 17	391	6.3	762	1	US-08-397-633A-26	Sequence 26, Appl
c 18	374.5	6.0	960	4	US-09-219-849-5	Sequence 5, Appl
c 19	371	6.0	960	4	US-09-219-849-6	Sequence 6, Appl
c 20	370.5	6.0	682	1	US-08-642-255-126	Sequence 126, App
c 21	370.5	6.0	682	1	US-08-397-633A-36	Sequence 36, Appl
c 22	365	5.7	761	2	US-08-707-237A-84	Sequence 84, Appl
c 23	365	5.7	762	1	US-08-642-255-114	Sequence 114, App
c 24	365	5.7	762	1	US-08-397-633A-26	Sequence 26, Appl
c 25	364.5	5.9	1078	3	US-08-963-825-21	Sequence 21, Appl
c 26	364.5	5.9	1078	4	US-09-500-811-21	Sequence 21, Appl
c 27	364.5	5.9	1078	4	US-09-570-573-21	Sequence 21, Appl
c 28	364.5	5.9	1078	4	US-09-548-608-21	Sequence 21, Appl
c 29	364	5.7	762	1	US-08-642-255-120	Sequence 120, App
c 30	364	5.7	762	1	US-08-397-633A-31	Sequence 31, Appl
c 31	360	5.8	1064	1	US-08-642-255-132	Sequence 62, Appl
c 32	357	5.7	829	1	US-08-642-255-132	Sequence 132, Appl
c 33	357	5.7	829	1	US-08-397-633A-53	Sequence 53, Appl
c 34	356.5	5.7	837	1	US-08-175-155-68	Sequence 103, Appl
c 35	356.5	5.7	837	1	US-08-477-509B-103	Sequence 103, Appl
c 36	356.5	5.7	837	1	US-08-642-255-101	Sequence 101, Appl
c 37	356.5	5.7	837	2	US-08-707-237A-75	Sequence 75, Appl
c 38	356.5	5.7	837	3	US-08-482-085B-103	Sequence 103, App
c 39	356.5	5.7	837	4	US-09-444-791A-103	Sequence 103, App
c 40	356.5	5.7	897	1	US-08-397-633A-50	Sequence 50, Appl
c 41	353.5	5.7	1065	1	US-08-642-255-72	Sequence 72, Appl
c 42	351.5	5.6	1057	3	US-08-931-820-4	Sequence 4, Appl
c 43	351	5.6	822	4	US-09-219-849-49	Sequence 49, Appl
c 44	345	5.5	1057	3	US-08-931-820-1	Sequence 1, Appl
c 45	344.5	5.5	633	1	US-08-642-255-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1  
US-09-020-956-113  
Sequence 113, Application US/09020956  
Patent No. 6261562  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS  
NUMBER OF SEQUENCES: 178  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/020,956  
FILING DATE: 09-FEB-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.427C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 113:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid



Tue Jun 10 11:28:25 2003

Best Local Similarity 78.9%; pred. NO. 2.3e-38;  
Matches 625; Conservative 0; Mismatches 128; Indels 39; Gaps 15;

QY	2	GAACACGCTGACGGCTGGCTCCGGGTGACAGCCCGCGCCTCGGCCAGGATCTGAG	61
Db	16	GAACACGCTGACGGCTGGCTCCGGGTGACAGCCCGCGCCTCGGCCAGGATCTGAG	75
QY	62	TGATGAGACGTGTCCCACTGAGGTGCCCCACAGCA-GCAGCTGTTGAGCATGGCGCTGAG	120
Db	76	TGATGAGACGTGTCCCACTGAGGTGCCCCACAGCATGCAGGTGTTGAGCATGGCGCTGAG	135
QY	121	AAGCTGGACC-GGCACCAAAAGGCGTGGCAGAAATGGCGCCGTG-GCTGATTCCTAGGCA-	177
Db	136	AAGCTGGACCTGGCACCAAAAGGCGTGGCAGAAATGGCGCCGTGCTGATTCCTAGGCAA	195
QY	178	----GTTGGCGGCAGCAGGAGAGGCCGCGCAGCTTCTGGAGCAGAGCCGAGAGCAAGC	233
Db	196	GTTCGGCGGAGCAGGAGATGAGGCCGCGCAGCTTCTTGGAGCAGAGCCGAGAGCAAGC	255
QY	234	AGTTCTGGAGTGCCCT---GAACGGCCCGCTGAGCCCTACCCGCC-TGGGCCCACTATGTC	289
Db	256	AGTTCTGGAGTGCCCTTGAACGGGCCCGCTGAGCCCTACCCGCCCTGCCCCACTATGTC	315
QY	290	CAGAGGCTGTGGGTGAGCGCGCTGCTCGGCGACCGGAAAGCCC--AGCTCTTTGCTGGTCA	347
Db	316	CAGAGGCTGTGGGTGAGCGCGCTGCTCGGCGACCGGAAAGCCCGCTCTTGTCTGGTCA	375
QY	348	ACCTGCTAAACCTTTGGCCTGGAGGTGTGTTTGGCGCAGCGCATCACTATGTGCCCGCTC	407
Db	376	ACCTGCTAAACCTTTGGCCTGGAGGTGTGTTTGGCGCAGCGCATCACTATGTGCCCGCTC	435
QY	408	TGC-----TGCTGGAAGTGGGGGTAGAGAGAAAGTTTCATGACCATGGTGTCTGGGCA-TTG	461
Db	436	TGCTTGCTTGGACACGTGGGGGTAGAGAGAAAGTTTCATGACCATGGTGTCTGGGCAATTG	495
QY	462	GTCCAGTGTGGCGCTGTGTGTCCCGCTCCTAGGCTCAGCCAGTGACC-ACGTGGCGT	520
Db	496	GTCCAGTGTGGCGCTGTGTGTCCCGCTCCTAGGCTCAAGCAGTGACCAACATGGGGT	555
QY	521	GGACGCTATGG-----CCGCGCGCGCCCTTCATCTGGGCACTGTCCCTTGGGCATCCGTG	574
Db	556	GGACGCTATGGCGCGCGCCCGCTTCATCTGGGCACTGTCCCTTGGGCATCCGTG	615
QY	575	CTGAGCCCTTTT-CTCATCCCAAGGCCC-----GGCTGGGTAGCAGGGGCTGCTGTGCC	626
Db	616	GTGAGCCTCTTTTCTTCATCCAGTGTGCGCGGTGGCTAAGCAGCGGCTTGGTGTGTC	675
QY	627	CGGATCCCAAGCCCTTGA-CCTGGCACTGCTATCTGGGCGTGGGCTGTGGACTTC	685
Db	676	TGATCCCAAGGCGCTGGGAGGCTAGGACTTGTCAATCCTGGGTGTGGGCTGTGCATCT	735
QY	686	TGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTGACCTTTCGGGACCCG	745
Db	736	TGTGGGCCAGGTTGTCCTCTCCCTTGSAGGCCGTGCT---TGECCTTTTCCGGGACCG	792
QY	746	GACCACTGTTCG	757
Db	793	GGCCTGTGTCG	804

RESULT	44
AWI35465/c	
LOCUS	
DEFINITION	AWI35465 370 bp mRNA linear EST 29-OCT-1999 UI-H-BII-aca-e-07-0-UI.s1 NCLCGAP_Sub3 Homo sapiens cDNA clone
ACCESSION	IMAGE:2713812 3', mRNA sequence.
VERSION	AWI35465
KEYWORDS	AWI35465.1 GI:6139598
SOURCE	EST. human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 370)

AUTHORS	JOURNAL
TITLE	COMMENT
...	...

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>,  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/LNL at:  
[www-bio1.lnl.gov/bbrp/image/image.html](http://www-bio1.lnl.gov/bbrp/image/image.html)  
Seq primer: M13 Forward  
POLYA=Yes.

FEATURES  
SOURCE

Location/Qualifiers

1. 370

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2713812"

/clone.lib="NCI\_CGAP\_Sub3"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pTTT3D-Pac (Pharmacia) with a modified polylinker; Site:1. Not I; Site:2. Eco RI; The NCI\_CGAP\_Sub3 library is a subtracted library derived from the NCI\_CGAP\_Sub1 library, which is a subtracted library derived from BI. BI constitutes a mixture of 21 normalized or subtracted NCI\_CGAP libraries: NCI\_CGAP\_Co10, NCI\_CGAP\_Pr22, NCI\_CGAP\_Pr28, NCI\_CGAP\_Co10, NCI\_CGAP\_Co16, NCI\_CGAP\_Kid5, NCI\_CGAP\_Kid12, NCI\_CGAP\_Kid3, NCI\_CGAP\_Kid11, NCI\_CGAP\_Lym2, NCI\_CGAP\_Pr2, NCI\_CGAP\_Co8, NCI\_CGAP\_CLL1, NCI\_CGAP\_Le12, NCI\_CGAP\_Pr23, NCI\_CGAP\_Lu5, NCI\_CGAP\_Lu24, NCI\_CGAP\_Lu19, NCI\_CGAP\_GC4, NCI\_CGAP\_GC6, NCI\_CGAP\_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:

NCI\_CGAP\_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clonoids 13232376-1323911, 14560008-1456775, 1500552-1502855); NCI\_CGAP\_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI\_CGAP\_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991, 1520904-1522439); NCI\_CGAP\_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI\_CGAP\_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids 985608-986759, 1101192-1101959, 1217928-1220615); NCI\_CGAP\_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described (Bonaldi, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.

TAG\_lib=NCI\_CGAP\_Pr22

TAG\_tissue=prostate

TAG\_seq=AAGTC"

	BASE COUNT	ORIGIN
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0
15	0	0
16	0	0
17	0	0
18	0	0
19	0	0
20	0	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	0	0
29	0	0
30	0	0
31	0	0
32	0	0
33	0	0
34	0	0
35	0	0
36	0	0
37	0	0
38	0	0
39	0	0
40	0	0
41	0	0
42	0	0
43	0	0
44	0	0
45	0	0
46	0	0
47	0	0
48	0	0
49	0	0
50	0	0
51	0	0
52	0	0
53	0	0
54	0	0
55	0	0
56	0	0
57	0	0
58	0	0
59	0	0
60	0	0
61	0	0
62	0	0
63	0	0
64	0	0
65	0	0
66	0	0
67	0	0
68	0	0
69	0	0
70	0	0
71	0	0
72	0	0
73	0	0
74	0	0
75	0	0
76	0	0
77	0	0
78	0	0
79	0	0
80	0	0
81	0	0
82	0	0
83	0	0
84	0	0
85	0	0
86	0	0
87	0	0
88	0	0
89	0	0
90	0	0
91	0	0
92	0	0
93	0	0
94	0	0
95	0	0
96	0	0
97	0	0
98	0	0
99	0	0
100	0	0

	Query Match	10.4%	Score 353;	DB 10;	Length 370;
	Best Local Similarity	100.0%;	Pred. No. 2.3e-37;		
	Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1991	CAGCTCCCGGTCCTGTATTAGCCCCCATGGGGCTGCCGGGCTGGCGGCACGTTTCTGTTCCT	2050		
Dd	370	CAGCTCCCGGTCCTGTATTAGCCCCCATGGGGCTGCCGGGCTGGCGGCACGTTTCTGTTCCT	311		
Qy	2051	GCCAAAGTAATATGGGCTCTCTGCTGGCACCCCTGTGCTGCTGAGTGCGTAGCTGCACAGC	2110		
Dd	310	GCCAAAGTAATATGGGCTCTCTGCTGGCACCCCTGTGCTGCTGAGTGCGTAGCTGCACAGC	251		







645 AGCTGGCACTGCTCATC 661  
 II III IIIII IIII  
 612 AGTAGGCCCTGCTGATC 628

RESULT 40  
 BB627844  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BB627844 650 bp mRNA linear EST 26-OCT-2001  
 BB627844 RIKEN full-length enriched, adult male urinary bladder Mus  
 musculus cDNA clone 9530042D02 5', mRNA sequence.  
 BB627844  
 BB627844.1 GI:16465416  
 EST.  
 house mouse.  
 Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;  
 1 (bases 1 to 650)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,  
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda  
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,  
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki  
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,  
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,  
 Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
 Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh  
 M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura  
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
 Hayashizaki,Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara  
 Y., and Hayashizaki,Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa  
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and  
 Hayashizaki,Y.  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.  
 e mouse tissues.  
 Location/Qualifiers  
 1. 650  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="9530042D02"  
 /clone\_lib="RIKEN full-length enriched, adult male urinary  
 bladder"  
 /sex="male"  
 /tissue\_type="urinary bladder"  
 /dev\_stage="adult"  
 /lab\_host="rn10a"

FEATURES  
 source





Query Match	12.08;	Score	409.4;	DB	10;	Length	692;
Best local Similarity	79.1%;	Pred. No.	6.5e-45;				
Matches	527;	Conservative	0;	Mismatches	127;	Indels	12;
Gaps	3;						
Qy	1	GGGAACGAGCTGCACGCGCTGGCTCGGTGTGACAGCGCGCGCTCGGCCAGGATCTGA	60				
Db							
Qy	37	GGGTCCGAGCTGCGACGCGCCAGCCCGAGGTGACGCCGACGCC	95				
Db							
Qy	61	GTGATGAGACGTGTCCCACTGAGGT--GCCCCACAGCAGCAGGTGTGAGCATGGGCTG	118				
Db							
Qy	96	CCGACGAGATGTGTCCCATCAAGCAAGGCACTAGATGTTGACGTGTTTACGCTGGGACG	155				
Db							
Qy	119	AGAAGCTGGACCGCACCAAGGGCTGGCAGAATGGCGCCTGGCTGATTCCTTAGGCAG	178				
Db							
Qy	156	AGATGCTGAAATTGGCATTAAGGGCTGGCAGAATGGGAACCTGGCTGCACCTAGGAGG	215				
Db							
Qy	179	TTGGCGCAGCAGGAGGAGGCCCGAGCTTCTGGAGCAGAGCCGAGACGACGATTC	238				
Db							
Qy	216	TTAGTGCTAGTGAGGAGGAGAAGCCAC-----GCCAGGGGTGACTCAAAGCAGCTG	266				
Db							
Qy	239	TGGAGTGCCTTGACGGGCCCCCTGAGCCCTACCGCCCTGGCCCACTATGGTCCAGAGGCTG	298				
Db							
Qy	267	TGGAGTATGTAGTAGCCCTTGGAAACCTACCTGCGCTGTCCATCATGATCCAGAGGCTG	326				
Db							
Qy	299	TGGTGTAGCGCCCTGCTGCGGCACCGGAAAGCCCCAGCTTGTGTGTCAACCTGCTAAACC	358				
Db							
Qy	327	TGGCCAGCCGTGCTGTACGCCACCGAAGCTCAGCTCCTGCTGGTCAACCTGCTCAACC	386				
Db							
Qy	359	TTTGGCCGTGAGGTGTGTTTGGCCGCAAGGCATACCTATGTGCCGCCCTGCTGCTGGAA	418				
Db							
Qy	387	TTTGGCCGTGAGGTGTGCCCTGGCTGCCGGGATACCTATGTGCCACCCCTTCTGCTGAA	446				
Db							
Qy	419	GTGGGGGTAGAGGAGATTCATGACCATGTTGCTGGGCATTTGTCAGGTGCTGGGCTG	478				
Db							
Qy	447	GTCGGGGTGGAGGAGAAATTCATGACCATGTTGTTGGGCATTTGCCCCAGTGTAGGCCTG	506				
Db							
Qy	479	GTCTGTGTCGCGTCTCTAGGCTCAGCCAGTGCACACCTGGGCTGGACGCTATGGCCGCCG	538				
Db							

```

VERSION BF789072.1 GI:12094108
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 759)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue procurement: Jeffrey F. Green, M.D.
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9810 row: 1 column: 07
High quality sequence start: 3
High quality sequence stop: 756.

FEATURES
    source
    1..759
        /organism="Mus musculus"
        /strain="FVB/N"
        /db_xref="taxon:10090"
        /clone="IMAGE:4222998"
        /clone_lib="NCI_CGAP_Kid14"
        /lab_host="DH10B (T1 phage-resistant)"
        /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
        Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
        Average insert size 1.75 kb. Constructed by Life
        Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT 131 a 218 c 243 g 167 t
ORIGIN
Query Match 12.3%; Score 419.2; DB 12; Length 759;
Best Local Similarity 79.4%; Pred. No. 3.2e-46;
Matches 575; Conservative 0; Mismatches 133; Indels 16; Gaps 6;

QY 9 GCGTCACGCGCTCGGTCGACAGCCGCGCTCGGCCAGGATCTGAGTGATGAG 68
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 16 GCTCGACGCGCCAGCCGCGGTCGACAGCCGCGCCGCGGCGGCGGCGGCGAG 74

QY 69 ACGTGTCCTCCACTGAGGT- -GCCCCACAGCAGCAGGTGTGAGCATGGCTGAGAAGCTG 126
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 75 ATGTGTCCCCATCAAGCAAGCAGCAGTATGTTGACGTGTTAGCGTGGGACGAGATGCT- 133

QY 127 GACCGGCACCAAGAGGCTGGCAGAAATGGGCGCTGCTGATTCCTAGGCGAGTTGGCGGC 186
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 134 GAATGGCACTAAGGGCTGGCAGAAATGGGAACCTGCTGCACCTAGGAGGTTAGTGCT 193

QY 187 ACAGAGGAGGAGGCGCCGACGCTTCTGGAGCAGAGCCGAGCAGGATCTTGGAGTGC 246
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 194 AGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 244

QY 247 CTGAAGGCGCCCTGAGCCCTTACCGCTGCGCCACTATGGTCAGAGGCTGTGGGTGAG 306
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 245 GTGAGTAGCCCTGGAACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304

QY 307 CCGCTGTGTCGGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 305 CCGTGTGCTACGCGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 364

QY 367 GGAGGTGTGTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 365 GGAGGTGTGTTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 424

QY 427 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 486
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 425 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484

```

```

QY 487 CCCGCTCCTAGGCTCAGCCAGTCACCACTGGCGTGGACGCTATGGCGCGCGGCCCTT 546
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 485 TCCACTCCTAGGCTCAGCCAGTCACCACTGGCGTGGCGTGGCGGCTATGGCGG- 543
QY 547 CATCTGGGCACTGCTCTGGGCACTCTGCTGGGCACTCTGCTGGGCACTCTCTTCTCATCCCAAGGCGCGGCTG 606
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 544 TATCTGGGCTTGTCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
QY 607 GCTAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 604 GCTGGCAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
QY 667 CGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
D 11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
D 664 AGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
QY 727 TGAC 730
D 11 | | |
D 722 CGAC 725

```

```

RESULT 34
BB627667
LOCUS BB627667
DEFINITION BB627667 RIKEN full-length enriched, adult male urinary bladder Mus
musculus cDNA clone 9530024M01 5', mRNA sequence.
BB627667
BB627667.1 GI:16465271
EST.
house mouse.
house musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, H., Okada,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system; 384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa,
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with

```















	AUTHORS	TITLE	JOURNAL	COMMENT
source	1. .491	/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="TWACF.1158206"		

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA library Preparation: Life Technologies, Inc.  
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

FEATURES SOURCE

BASE COUNT  
ORIGIN

RESULT	25
LOCUS	Bf785813
DEFINITION	Bf785813 929 bp mRNA linear EST 12-JAN-2001 602112437F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4240617
ACCESSION	Bf785813
VERSION	Bf785813.1 GI:12090849
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 929)

```

QY 965 CCACGAGAGGGCTGTGCGGCCCTCTTGTGCGCCCACTGCTGTCCATGCCGGCCGG 1024
Db 302 CCACGAGAGGGCTGTGCGGCCCTCTTGTGCGCCCACTGCTGTCCATGCCGGCCGG 361
QY 1025 TTGGCTTTCGGAACTGGGCGGCTCTTCCGCGGCTGCACAGCTGTGCTGCCGATG 1084
Db 362 TTGGCTTTCGGAACTGGGCGGCTCTTCCGCGGCTGCACAGCTGTGCTGCCGATG 421
QY 1085 CCGCGCACTTGGCGGCTCTTGGTGGCTGAGCTGTGCGAGTGGATGGCACTCATGACC 1144
Db 422 CCGCGCACTTGGCGGCTCTTGGTGGCTGAGCTGTGCGAGTGGATGGCACTCATGACC 481
QY 1145 T 1145
Db 482 T 482

RESULT 23
BG173136
LOCUS 602335411F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4458602 5',
DEFINITION mRNA sequence.
ACCESSION BG173136
VERSION BG173136.1 GI:12679748
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 969)
JOURNAL NTH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL10257 row: e column: 03
High quality sequence stop: 608.
Location/Qualifiers
1. 969
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4458602"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 199 a 289 c 287 g 194 t
ORIGIN

Query Match 13.5%; Score 459; DB 12; Length 969;
Best Local Similarity 82.9%; Pred. No. 1.6e-51;
Matches 583; Conservative 0; Mismatches 115; Indels 5; Gaps 5;

QY 839 AGTGCCCTGGCCCTTACCTGGGACACCGAGGAGTGCCTTTGGCTGCTCACCCTC 898
Db 1 AGGTCTTGGCCCTTACCTGGGACTAGGAGATGCCTTTGGCTCCTCACCCTC 60
QY 899 ATCTCTCCTCCTGCTAGCAGCCACACTGCTGGTGGCTGAGGAGGAGCGCTGGGCCC 958
Db 61 ATTTCTCTCATCTGCTGCGAGCCACTCTGTTTGTGACGAGGAGGAGTACTGGGCCCA 120

```

```

QY 959 ACCGAGCCAGAGAGGGCTGTGCGGCCCTCTTGTGCGCCCACTGCTGTCCATGCCGG 1018
Db 121 CCGGAGCCCGCAGAGGGTGTGTTGGTCTCTGCGGTGTGCGCCGATGTGCCCATGCCAC 180
QY 1019 GCCCGCTTGGCTTTCCGGAACCTGGGCGGCTGCTTCCCGGCTGCACAGCTGTGCTGC 1078
Db 181 GTTGGCTTGGCTTTCCGGAATCTGGTACCCTGTTCCTCCCGGCTGCAGAGCTGTGCTGC 240
QY 1079 CGCATGCCCGCCGACCCCTGCGCGGCTCTTCGTTGGTGTGAGCTGTGAGTGGATGCATC 1138
Db 241 CGCATGCCCTGCGACCCCTACGCGGCTCTTGTGGCTGAGCTGTGAGCTGGATGCATC 300
QY 1139 ATGACCTTCCACGCTGTTTACACGAGTTCGTTGGCGGAGGGGCTGTACAGGCGCTGCC 1198
Db 301 ATGACTTTCACACTGTTTACACGAGTTCGTTGGGAGAGGGGCTGTACCA -GGTGTACCC 359
QY 1199 AGAGCTGAGCCGCGGACCGGAGGAGGAGGAGGCTTTCGATGGGAGGAGGAGGAGG 1258
Db 360 AGAGCCGAGCCAGGACCGAGG -CCGAGAGACTATGATGAGGCAATTTCGAATGGCAGC 418
QY 1259 CTGGGGCTTCTCTGAGTGGCCCATCTCCCTGGTCTTCTCTGCTGTGATGACCGGCTG 1318
Db 419 CTGGGGCTTCTCTGAGTGGCCCATCTCCCTGGTCTTCTCTGCTGTGATGACAGGCTG 478
QY 1319 GTGACGAGTTCGGCACTCGAGCACTTATTTGGCCAGTGTGGCAGCTTTCCTCTGGCT 1378
Db 479 GTACAGAGTTCGGGACACAGGCTCAGTCTATCTGGCCAGTGTGATGAGCTTTCCTGTGGCT 538
QY 1379 GCCGTGTCACATGCTCTCCACAGTGTGGCCGTGTGACAGCTTCAGCGCCCTCACC 1438
Db 539 GCCGTGTCACCTG -CTGTCACACAGCTGTGGTGTGAGTACAGC -TCAGGTGCCCTCACC 596
QY 1439 GGGTTCACCTTCTCAGCCCTCGAGTCTGCTCCCTACACACTGGCTCTCCTTACACCGG 1498
Db 597 GGGTTCACCTTCTCAGCCCTCGAGTCTGCTCTTAAAGA -TCGCTCTCTTACACCGG 655
QY 1499 GAGAAAGAGGTGTTCTCTGCCCAATACCGAGGAGGAGGAGGAGGAGGAGGAG 1541
Db 656 AGAAAGCCACTTGTTCGCGCCCAATACCGAGGAGGAGGAGGAGGAGGAGGAGGAG 698

RESULT 24
AA631143
LOCUS ng75908.s1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1158206 3',
DEFINITION mRNA sequence.
ACCESSION AA631143
VERSION AA631143.1 GI:2553754
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 491)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 919 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 472.
Location/Qualifiers

```

FEATURES

found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM11855 row: p column: 02  
 High quality sequence stop: 778.  
 Location/Qualifiers

# FEATURES

source

1..901  
 /organism="Mus musculus"  
 /strain="129,C57BL/6J,FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5337073"  
 /clone\_lib="NCL CGAP Mam3"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="10 months"  
 /lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;  
 Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
 Reference for transgenic model: Xu et al., Nature Genetics  
 22, 37-43 (1999)."

BASE COUNT 151 a 260 c 288 g 202 t

## ORIGIN

Query Match 13.6%; Score 463; DB 13; Length 901;  
 Best Local Similarity 78.5%; Pred. No. 5e-52;  
 Matches 684; Conservative 0; Mismatches 165; Indels 22; Gaps 10;

QY 1 GGGAACACGCTGCACGGCTGGCTCGGGTGACAGCGCGCGCTCGCCAGGATCTGA 60  
 DB 42 GTGTCCGAGCTGCACGGCCGACCGCCAGGTGACAGCGCGCGCTCGCCAGGATCTGA 100  
 QY 61 GTGATGAGACGTGCTCCCACTAGGT---GCCCCACAGCAGCAGGTGTGACATGGCT 117  
 DB 101 CCGACGAGATGTCTCCCACTAGGT---GCCCCACAGCAGCAGGTGTGACATGGCT 160  
 QY 118 GAGAAGCTGGACCGCGCACCAAGAGGCTGGCAAAATGGCGCTGGCTGATTCTTAGCA 177  
 DB 161 GAGATGCTGAATGGCACTAAGAGGCTGGCAAAATGGCAAAATGGCAAAATGGCA 220  
 QY 178 GTTGGCGGACGACGAGAGAGGCGGAGCTTCTGGAGCAGAGCGGAGCAAGAGCT 237  
 DB 221 GTTAGTGTCTAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 271  
 QY 238 CTGGAGTGTCTGAACGGCCCTGAGCCCTACCTCCCGCTGGCCACTATGTTCCAGAGCT 297  
 DB 272 GTGGAGTATGAGTAGCCCTGGAACTCTAGCTCCCTGTCCCTATGATGATCCAGAGCT 331  
 QY 298 GTGGGTGAGCCGCTGCTGCGCACCGGAAAGCCAGCTCTGTGCTCAACCTGCTAAC 357  
 DB 332 GTGGGCGACCGCTGCTGCTACGGCACCGGAAAGCTCAGCTCCTGTGTTCAACCTGCTAC 391  
 QY 358 CTTTGGCCTGAGAGTGTGTTGGCGCAGGCAATACCTATGTGGCGCTCTGTGCTGGA 417  
 DB 392 CTTTGGCCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451  
 QY 418 AGTGGGGGTAGAGGAGAGTTCATGACCATGCTGTGGCATTTGCTCCAGTGTGGGCT 477  
 DB 452 AGTGGGGGTAGAGGAGAGTTCATGACCATGCTGTGGCATTTGCTCCAGTGTGGGCT 511  
 QY 478 GGTCTGTGCTCCGCTCTTAGGCTCAGCCAGTGAACACTGGGCTGGACGCTATGGCCGCG 537  
 DB 512 GGTCTGTGCTCCGCTCTTAGGCTCAGCCAGTGAACACTGGGCTGGACGCTATGGCCGCG 571  
 QY 538 CCGGCGCTTCATCTGGGCTGCTGCTGGCATCTGCTGAGCTCTTCTTCATCCCAAG 597  
 DB 572 GAGACCTTATCTGGGCTGCTGCTGGCATCTGCTGAGCTCTTCTTCATCCCAAG 631  
 QY 598 GGCCTGCTGCTAGGAGGCTGCTGCTGGCCGAGTCCAGGCGCTTGGAGCTGCTGCTGCT 657  
 DB 632 GGCCTGCTGCTAGGAGGCTGCTGCTGGCCGAGTCCAGGCGCTTGGAGCTGCTGCTGCT 691  
 QY 658 CATCTGGCGGTGGGCTGCTGAGACTTCTGTGGCGAGGTGCT--TCACCTCCACTGGAG 715

DB 692 GATCTTGGAGTGGGGCTGCTGGACGT-TGTGGCAGCTGTGCTTGTACTTCCATTGGAG 750  
 QY 716 GCCCTGCTCTGA--CCTCTTCGGGACCGGACAC-TGTGCCAGGCTTACTCTGTC 772  
 DB 751 GCCTTACTCTCCGAACCTCTTCCGGGACCCAGACCACTTGGCGCAAGGCTTCTCTGTC 810  
 QY 773 -TATGCTTTCATGATCAGTCTTGGGGCTGCCT--GGCTACCTCCTCCTGCCATTGACT 830  
 DB 811 TTACGCTTCTCTGATCAGCTTGGGGCTGCCTGGGGTACCCTTCTTACTGCGCGCTT 870  
 QY 831 GGGACA--CAGTGGCCCTGGCCCTACTCTGG 860  
 DB 871 GGGACACCCAGCGTCTTGGCCCTACTGGG 901

## RESULT 22

AF109299

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

1..482

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IPCA-2"

/clone\_lib="Homo sapiens prostate adult"

/tissue\_type="prostate"

/dev\_stage="adult"

/note="multiple clone assembly from multiple libraries and donors"

BASE COUNT 57 a 179 c 141 g 105 t

## ORIGIN

Query Match 13.6%; Score 462.6; DB 9; Length 482;

Best Local Similarity 99.0%; Pred. No. 7.6e-52;

Matches 476; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 666 GCGTGGGGCTGTGGACTTCTGT--GGCCAGTGTGCTTCACTCCACTGGAGGCCCTGCTC 724

DB 2 GCGTGGGGCTGTGGACTTCTGTGGCCAGTGTGCTTCACTCCACTGGAGGCCCTGCTC 61

QY 725 TCTGACCTTCTCCGGACCCGACCACTGTGCGCAGGCGCTTCTGTCTATGCTTTCATG 784

DB 62 TCTGACCTTCTCCGGACCCGACCACTGTGCGCAGGCGCTTCTGTCTATGCTTTCATG 121

QY 785 ATGAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCGCATTTGACTGGACACCACTGCC 844

DB 122 ATTAGTCTTGGGGCTGCTGGGCTACCTCTGCTGCGCATTTGACTGGACACCACTGCC 181

QY 845 CTGGCCCTTACCTGGGACCCAGGAGGAGTGTCTTTGGCTGTCTACCCCTCATCTTC 904

DB 182 CTGGCCCTTACCTGGGACCCAGGAGGAGTGTCTTTGGCTGTCTACCCCTCATCTTC 241

QY 905 CTCACCTGCTAGACCCACACTGCTGTGTGGCTGTGAGGAGCAGCGTGGGCCCCACCGAG 964

DB 242 CTCACCTGCTAGACCCACACTGCTGTGTGGCTGTGAGGAGCAGCGTGGGCCCCACCGAG 301

QY 546 TCATCTGGGCACTGCTCTTGGGCATCTCTGAGCCTCTTTCTCATCCCAAGGGCGGCT 605  
 D 542 TTATCTGGGCTTGTCCCTGGGTGCTCTGCTAGCCTCTTTCTCATCCCGAGGGCTGGCT 601  
 QY 606 GGCTAGCAGGCTGCTGTGTCGCCGGATCCAGGCCCCGAGCTGGGCACTGCTCATCTCTGG 665  
 D 602 GGCTGAGGCACTGCTGT-ACCAGACACAGGCCCCGAGTGGCCCTGCTGATCTTGG 660  
 QY 666 GCCTGGGGCTGCTGGCACTTCTG--TGGCCAGGTGCTCTCACTCCACTGGAGCCCTGCT 723  
 D 661 GAGTGGGGCTGCTGGCACTTTTGTGGCCAAAGGTGCTTTACTCCATTGGAGGCCCTTACT 720  
 QY 724 CTCTGACCTTCTCCGGGACCGGACCACTGTGCCAGGCTCTACTC-TGTCTATGCTTCA 782  
 D 721 CTCGAGCTCTCCGGGAGCAGCAACATTGGCGCAAGGCTTCTCTGTCTACGCTTCA 780  
 QY 783 TGATCAGTCTTGGGGCTGCTGGGTACTCTCTGCTGCCATTGACTGGGACACAGTG 842  
 D 781 TGATCAGCTTGGGGCTGCTGGGTACTCTTTTACCTGGCATGACTGGAACACAAG-G 839  
 QY 843 CCCTGGCCCTACCTGGGCAACCCAGG 869  
 D 840 TGCTGGGCCCCCTACCTGGGTAATCAGG 866

RESULT 20  
 LOCUS BG174399 786 bp mRNA linear EST 06-FEB-2001  
 DEFINITION 602334219F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:4457452 5',  
 mRNA sequence.

ACCESSION BG174399

VERSION BG174399.1 GI:12681102

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 786)

NIH-MGC <http://mgi.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM10254 row: e column: 05

High quality sequence stop: 666.

Location/Qualifiers

1. .786

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4457452"

/clone\_lib="NCI\_CGAP\_Mam1"

/tissue\_type="tumor, biopsy sample"

/dev\_stage="10 months, virgin"

/lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 121 a 246 c 234 g 185 t

ORIGIN

Query Match

Best Local Similarity 14.7%; Score 499.6; DB 12; Length 786;

Matches 579; Conservative 0; Mismatches 94; Indels 3; Gaps 2;

QY 593 CCAAGCGCGGCTGGCTAGCAGGGCTGTGTCGCCGGATCCAGGCCCTCGAGCTGC 652  
 D 1 CCGAGGGCTGGCTGGCTGGCAGGACTGCTGTACCCAGAGACAGGCCCTCGAGTAGGCC 60  
 QY 653 CTGCTCATCTGCTGGGGCTGGGGCTGCTGACCTTCTGTGCCAGGTGTGCTTCACTCCACTG 712  
 D 61 CTGCTGATCTTGGGAGTGGGGCTGCTGACCTTTTGTGCCAGGTGTGCTTACTCCATTTG 120  
 QY 713 GAGGCCCTGCTCTCTGACCTCTTCCGGGACCGGACCACTGTGCCAGGCGCTACTCT-ET 771  
 D 121 GAGGCCCTTACTCTCTCGACCTCTTCCGGGACCGGACCACTGTGCCAGGCGCTACTCTCTCTGG 180  
 QY 772 CTATGCCCTTCATGATGACGTCTTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTG 831  
 D 181 GTACGCCCTTCATGATGACGTCTTGGGGCTGCTGGGCTACCTCTTACCTGCCATGACTG 240  
 QY 832 GGACACCACTGCTGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891  
 D 241 GGACACCACTGCTGCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 892 CACCGCTCATCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951  
 D 301 CACCGCTCATCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 QY 952 GGGCCCCACCGAGCCAGAGGGCTGCTGCGGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1011  
 D 361 GGGCCCCACCGAGCCAGAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 QY 1012 ATGCCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1071  
 D 421 ATGCCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 1072 GTGCT 1131  
 D 481 GTGCT 540  
 QY 1132 GGCACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1191  
 D 541 GGCACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 1192 CGTGGCCAGAGCTGAGCGGGGCGGACCGAGCCCGGAGACACTATGATGAGGCGTTCCGA- 1250  
 D 601 TGATACCCAGAGCGGAGCGGAGCGGAGCCCGGAGACACTATGATGAGGCGTTCCGA 660  
 QY 1251 -TGGGCAGCTGGGGC 1265  
 D 661 TGGGCAGCTGGGGC 676

RESULT 21

LOCUS BI650119

DEFINITION

BI650119

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI650119 901 bp mRNA linear EST 12-SEP-2001  
 603296208F1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5337073 5',  
 mRNA sequence.

BI650119

BI650119.1 GI:15564355

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 901)

NIH-MGC <http://mgi.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be



/clone="IMAGE:5473696"  
 /tissue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pOTB7; Site:1; XhoI: Site:2;  
 EcoRI: cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GCACAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."  
 BASE COUNT 310 a 665 c 258 g 418 t 16 others  
 ORIGIN

Query Match 15.0%; Score 512.8; DB 14; Length 1667;  
 Best Local Similarity 97.4%; Pred. No. 1e-58;  
 Matches 553; Conservative 0; Mismatches 12; Indels 3; Gaps 3;  
 QY 1802 TCCTGTTATGGGCTCCATTGTCAGCTCAGCAGTCTGTCTACCTGCCTATATGTGTCT 1861  
 Db 1 TCCTGTTATGGGCTCCATTGTCAGCTCAGCAGTCTGTCTACCTGCCTATATGTGTCT 60  
 QY 1862 GCCGAGGCTGGTCTGTGCGCATTTACTTTGTCTACAGGTAGTATTTACAGAGC 1921  
 Db 61 GCCGAGGCTGGTCTGTGCGCATTTACTTTGTCTACAGGTAGTATTTACAGAGC 120  
 QY 1922 GACTTGGCCAAATACTACGCTAGAAACTTCCAGCACATTTGGGTGGAGGCCCTGCCTC 1981  
 Db 121 GACTTGGCCAAATACTACGCTAGAAACTTCCAGCACATTTGGGTGGAGGCCCTGCCTC 180  
 QY 1982 ACTGGGTCCAGCTCCCGCTCTGTAGTACCCCATGGGCTCCCGGCTGGCGCCAGTT 2041  
 Db 181 ACTGGGTCCAGCTCCCGCTCTGTAGTACCCCATGGGCTCCCGGCTGGCGCCAGTT 240  
 QY 2042 TCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCTCTGCTGTAGTGGCTGAG 2101  
 Db 241 TCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCTCTGCTGTAGTGGCTGAG 300  
 QY 2102 CTGCACAGCTGGGGCTGGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2161  
 Db 301 CTGCACAGCTGGGGCTGGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
 QY 2162 CTGGAGGCTCTCCAGGGGTTTCACTCGAGTATACCCAGGCTCAGGGCTCAGAGGCTCA 2221  
 Db 361 CTGGAGGCTCTCCAGGGGTTTCACTCGAGTATACCCAGGCTCAGGGCTCAGAGGCTCA 420  
 QY 2222 TGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGTAG-C 2280  
 Db 421 TGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGTAGCC 480  
 QY 2281 CTCCTAGTTGAGACACACCTA-GAGAAGGTTTGGAGCTGAATAAAGTCAAGTCACT 2339  
 Db 481 CTCCTAGTTGAGACACACCTAAGAGGCTCTTGGAGCTGAATAAAGTCAAGTCACT 540  
 QY 2340 GG-TTTCCCATCTTAAGCCCTTAACC 2366  
 Db 541 GGTTTCCCATCTTAAGCCCTTAACC 568

RESULT 19  
 BG864609  
 LOCUS  
 DEFINITION 872 bp mRNA linear EST 29-MAY-2001  
 602798469F1 NCI\_CGAP\_Mam4 Mus musculus cDNA clone IMAGE:4919513 5',  
 mRNA sequence.  
 ACCESSION BG864609  
 VERSION BG864609.1 GI:14215147  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 872)

# AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth  
 Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM10834 row: i column: 18  
 High quality sequence stop: 738.

## FEATURES Source

1..872  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="NMRI"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4919513"  
 /clone\_lib="NCI\_CGAP\_Mam4"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; SalI:  
 Site:2; NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies, Inc. Investigators  
 providing samples: Lothar Hennighausen/Priscilla Furth,  
 NIH Reference for transgenic model: Li et al., Cell Growth  
 and Differentiation 7, 3-11 (1996)."  
 BASE COUNT 157 a 242 c 276 g 197 t  
 ORIGIN

Query Match 14.9%; Score 508.6; DB 12; Length 872;  
 Best Local Similarity 79.7%; Pred. No. 5.1e-58;  
 Matches 691; Conservative 0; Mismatches 159; Indels 17; Gaps 7;  
 QY 8 AGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCTCGGCCAGGATCTGAGTGATGA 67  
 Db 12 AGCTCGCAGCGCGCCAGCCAGGTGACAGCCGCGCGCTCGGCCAGGATCTGAGTGATGA 70  
 QY 68 GACGTGTCCCTACTAGGT--GCCCCACAGCAGCAGGTGTGACGATGGCTGAGAAAGCT 125  
 Db 71 GATGTGTCCCATCAAGCAAGGCACTAGATGGTGTGACGTGTAGCGTGGGACGAGATGCT 130  
 QY 126 GGACCGGCACCAAGGGCTGGCAGAAATGGCGCTGCTGCTGATTCCTAGCAGTGGCGG 185  
 Db 131 GAATGTGCACTAAAGGCTGGCAGAAATGGCACTGGCTGCACCTTAGAGGTTAGTGC 190  
 QY 186 CAGCAAGGAGGAGGCGCGCAGCTTCTGGAGCAGAGCCGAGAGCAAGCAGTCTTGGAGTG 245  
 Db 191 TAGTGAGGAGGAGAGGCGC-----GGCAGGGCTGACTCAAGCAGCTGTGGAGTA 241  
 QY 246 CCTCAAGCGCCCTGAGCCCTACCCGCTTACCCGCTTGGCCCATATGGTCCAGAGCTGTGGTGA 305  
 Db 242 TGTGAGTAGCCCTTGGAAACCTTACCTGCTTCCATCATGATCCAGAGCTGTGGGCCA 301  
 QY 306 GCCGCTGTGCGGCACCGGAAAGCCAGCTTGTGCTGGTCAACCTGCTAACCTTTGGCC 365  
 Db 302 GCCGCTGTGAGCAGCAGGAAAGCTCAGCTCTGCTGGTCAACCTGCTAACCTTTGGCC 361  
 QY 366 TGGAGGTGTGTTTGGCGCAGGCAATACCTATGTGCGCCCTCTGCTGCTGGAAGTGGGG 425  
 Db 362 TGGAGGTGTGCTGGCTGCCGGCATTACCTATGTGCGCAGCCCTTCTGCTGGAAGTGGGG 421  
 QY 426 TAGAGGAGAGTTCATGACCATGGTGTGGCATTGGTCCAGTGTGGGCTGGTCTGTG 485  
 Db 422 TGGAGGAGAAATTCATGACCATGGTGTGGCATTGGCCAGTGTGGCTGTGGTCTG 481  
 QY 486 TCCGCTCTCTAGGCTCAGCCAGTGGCCACTTGGCTGGCAGCTATGGCGCGCGCCGCCCT 545  
 Db 482 TTCACCTCTAGGCTCAGCCAGTGGCCAGTGGCGCTGGGCGCTATGGCCCGCGGAGACCT 541

```

Db      302 CTAATATTGGTGGGATCCCAACAATCAGGTCCCTGAGATAGTGGTGCAT 361
QY      TGGGCTGATATGCCAGAACTTCTCTCTCTGGGGTCTGGCCGCCCAAAATGCTAAAC 2783
Db      362 TGGGCTGATATGCCAGAACTTCTCTCTCTGGGGTCTGGCCGCCCAAAATGCTAAAC 421
QY      2784 CAGGACCTTGGAAATCTTACTCATCCCAATGATAATTCACCAAGTTT 2843
Db      422 CAGGACCTTGGAAATCTTACTCATCCCAATGATAATTCACCAAGTTT 481
QY      2844 AGGCTGTTGAA-GGAAGTAGAGGTGGGGTTCAGGTCTCAAGGGTTCCTT-AACCAC 2901
Db      482 AGGCTGTTGAAAGGAGTAGAGGTGGGGTTCAGGTCTCAAGGGTTCCTTAAACCAC 541
QY      2902 CCCTCTCTCTGTCACCAAGCTT-GTTCCTCCCTCCCTTCCACTCCCTC 2948
Db      542 CCCTCTCTCTGTCACCAAGCTTGGGTTCCTCCCTTCCACTCCCTC 589

RESULT 17
LOCUS   BG122427
DEFINITION 602353324f1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4451572 5',
mRNA sequence.
ACCESSION BG122427
VERSION   BG122427.1 GI:12615936
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 715)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10238 row: p column: 05
High quality sequence stop: 689.
Location/Qualifiers
1..715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4451572"
/clone.lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
Note="Organ: Liver; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 176 a 174 c 175 g 190 t
ORIGIN

Query Match 15.4%; Score 525; DB 12; Length 715;
Best Local Similarity 94.5%; Pred. No. 3.9e-60;
Matches 664; Conservative 0; Mismatches 25; Indels 14; Gaps 11;

QY      2211 GAAGGGCTCATGCTGAGTGGGGTCTCGAGTGGATTACCCAGGCTCAGGGTT 2270
Db      1 GAAGGGCTCCATGCTGAGTGGGGTCTCGAGTGGATTACCCAGGCTCAGGGTT 60
QY      2271 AACAGCTAGCTCTCTAGTTGAGACACACCTAGAGAGGGTTTGGGAGCTGAATAAAT 2330
Db      61 AACAGCTAGCTCTCTAGTTGAGACACACCTAGAGAGGGT-TTTTGGGAGCTGAATAAAT 119

2331 CAGTCACCTGGTTTCCCATCTCTAAGCCCTTAACTGCAGCTTCGTTAAATAGCTCT 2390
120 CAGTCACCTGGTTTCCCATCTCTAAGCCCTTAACTGCAGCTTCGTTAAATAGCTCT 179
2391 TGATGGGAGTTTCTAGGATGAACACTCCTCCATGGGATTTGAACATATG-ACCTATT 2449
180 TGATGGGAGTTTCTAGGATGAACACTCCTCCATGGGATTTGAACATATGAAAGTTAT 239
2450 GTAGGGGAAGAGTCTGAGGGCAACACACAGACAGGTCCTCCCTCAGCCACAGCACT 2509
240 GTAGGGGAAGAGTCTGAGGGCAACACACAGACAGGTCCTCCCTCAG-CCACAGCACT 298
2510 GTCTTTTGTGATCCACCCCTCTTACCTTTTATCAGATGTGGCTGTGGTCTCTTC 2569
299 GTC-TTTTGTGATCCACCCCTCTTACCTTTTATCAGATGTGGCTGTGGCTCTTC 356
2570 TGTGTCATCAGACAGACAGCACTTTAAATATTTAACTTTATTTAAACAAGTAGA 2629
357 TGTGTCATCAGACAGACAGCACTTTAAATATTTAACTTTATTTAAACAAGTAGA 416
2630 AGGGAATCCATCTAGCTTTTCTGTGGTGTCTAATATTTGGGTAGG-TGGGGGAT 2688
417 AGGGAATCCATCTAGCTTTTCTGTGGTGTCTAATATTTGGGTAGG-TGGGGGAT 476
2689 CCCCAACAATCAGTCCCTTGAGATAGCTGGTCAATTTGGGCTGATCATGCCAGAATC-TT 2747
477 CCCCAACAATCAGTCCCTTGAGATAGCTGGTCAATTTGGGCTGATCATGCCAGAATC-TT 536
2748 CTCTCTCTGGGCTCTGGCCCCCAAAATGCTTAACCCAGGACCTTGAAATTTCTACTCAT 2807
537 CTCTCTCTGGGCTCTGGCCCC--AAATGCTTAACCCAGGA-CTGGGAATTTCTACTCAT 593
2808 CCCAAATGATAATTCAAATGCTGTACCCAGGTTAGGTTAGGGAAGGTAGAGGG 2867
594 GCCAAATGATAATTCAAATGCTGTACCC--AGGTAGGTTAGGGAAGGTAGAGGG 651
2868 TGGGGCTTCAAGTCTCAACGGTTCCTTAACCCACCTCTCTCT 2910
652 GGGGCTTCAGGGTCTCAACGG--TACCTAACCCACCTCTCTCT 692

RESULT 18
LOCUS   BM912193
DEFINITION 6613122 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5473696
5', mRNA sequence.
ACCESSION BM912193
VERSION   BM912193.1 GI:19362572
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1667)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: DICMI985 row: 1 column: 17
High quality sequence stop: 448.
Location/Qualifiers
1..1667
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

```

High quality sequence start: 3  
 High quality sequence stop: 602.  
 Location/Qualifiers  
 1..1035  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4905674"  
 /clone\_lib="NIH\_MGC\_17"  
 /tissue\_type="rhabdomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: muscle; Vector: pORF7; Site\_1: EcoRI; Site\_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 227 a 273 c 341 g 194 t  
 ORIGIN

Query Match 15.8%; Score 538.2; DB 12; Length 1035;  
 Best Local Similarity 93.1%; Pred. No. 5.9e-62;  
 Matches 632; Conservative 0; Mismatches 33; Indels 14; Gaps 6;  
 QY 7 CAGCCTGCACGCGTGGCTCGGCTGACAGCGCGCGCTCGGCCAGGATCTCAGTGATG 66  
 Db |||||||  
 QY 2 CAGCCTGCACGCGTGGCTCGGCTGACAGCGCGCGCTCGGCCAGGATCTCAGTGATG 61  
 Db |||||||  
 QY 67 AGACGTGTCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGTGAGAGCTG 126  
 Db |||||||  
 QY 62 AGACGTGTCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGTGAGAGCTG 121  
 Db |||||||  
 QY 127 GACCGGCACCAAGAGGTGGCAGAAATGGCGCTGGCTGATCTTCTAGGCAGTTGGCGGC 186  
 Db |||||||  
 QY 122 GACCGGCACCAAGAGGTGGCAGAAATGGCGCTGGCTGATCTTCTAGGCAGTTGGCGGC 181  
 Db |||||||  
 QY 187 AGCAAGAGGAGAGCGCCAGCTTCTTGAGCAGAGCCGAGCAGAGCAGTTCTGGAGTGC 246  
 Db |||||||  
 QY 182 AGCAAGAGGAGAGCGCCAGCTTCTTGAGCAGAGCCGAGCAGAGCAGTTCTGGAGTGC 241  
 Db |||||||  
 QY 247 CTGACAGCGCCCTGAGCCCTACCGCTGCGCCACTATGTCAGAGGCTGTGGGTGAG 306  
 Db |||||||  
 QY 242 CTGACAGCGCCCTGAGCCCTACCGCTGCGCCACTATGTCAGAGGCTGTGGGTGAG 301  
 Db |||||||  
 QY 307 CGCCTGCTGGCGGACCGGAAAGCCAGCTCTTGTGCTCAACCTGCTAACCTTTGGCCT 366  
 Db |||||||  
 QY 302 CGCCTGCTGGCGGACCGGAAAGCCAGCTCTTGTGCTCAACCTGCTAACCTTTGGCCT 361  
 Db |||||||  
 QY 367 GGAGTGTGTTGGCGGAGGATCACCCTATGTCGGCTCTGCTGCTGGAAGTGGGGT 426  
 Db |||||||  
 QY 362 GGAGTGTGTTGGCGGAGGATCACCCTATGTCGGCTCTGCTGCTGGAAGTGGGGT 421  
 Db |||||||  
 QY 427 AGAGGAGAAGTTCATGACCATGGTGTGGCATTTGGTCACTGCTGGCCCTGCTGCTGT 486  
 Db |||||||  
 QY 422 AGAGGAGAAGTTCATGACCATGGTGTGGCATTTGGTCACTGCTGGCCCTGCTGCTGT 481  
 Db |||||||  
 QY 487 CCGCG-TCCTAGGTCA-GCCAGTGACCACTGGCTGGACGCTATGGCCGCGCGCGCC 544  
 Db |||||||  
 QY 482 CCGCATCTAGGCTCAGCCAGTGACCACTGGGCTGGACGCTATGGCCGCGCGCGCC 541  
 Db |||||||  
 QY 545 TTCATCTGGGCACTGCTCTGGCATCTGCTGA-----GCTCTTCTCATCCCAAGG 598  
 Db |||||||  
 QY 542 TTCATCTGGGCACTGCTCTGGCATCTGCTGA-----GCTCTTCTCATCCCAAGG 601  
 Db |||||||  
 QY 599 GCGGCTGGTACGAC----GGCTGCTGTCGCC-GGATCCAGCGCCCTGGAGCTGGCAC 653  
 Db |||||||  
 QY 602 GCGGCTGGTACGACGACTGCTGCTGGCAGGATACAGGACCTTGGAGCTGG-AC 660  
 Db |||||||  
 QY 654 TGCTCATCTGGGCGTGGG 672  
 Db |||||||  
 QY 661 TGCTCATCTGGGCGTGGG 679  
 Db |||||||

RESULT 16  
 BG469586  
 LOCUS  
 DEFINITION  
 602533622F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4661500 5',  
 mRNA sequence.  
 ACCESSION  
 BG469586  
 VERSION  
 BG469586.1 GI:13401861  
 KEYWORDS  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 589)  
 NIH-MGC http://mgc.nci.nih.gov/  
 AUTHORS  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE  
 Unpublished (1999)  
 JOURNAL  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabps-re@mail.nih.gov  
 COMMENT  
 Tissue procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: NIH Intramural Sequencing Center  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 plate: LLCM1460 row: k column: 05  
 High quality sequence stop: 587.  
 Location/Qualifiers

FEATURES  
 source

1..589  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4661500"  
 /clone\_lib="NIH\_MGC\_15"  
 /tissue\_type="adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: colon; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 135 a 153 c 133 g 168 t  
 ORIGIN

Query Match 15.7%; Score 534.8; DB 12; Length 589;  
 Best Local Similarity 98.8%; Pred. No. 2.2e-61;  
 Matches 581; Conservative 0; Mismatches 2; Indels 5; Gaps 4;  
 QY 2366 CTCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTCTTAGGATGAACACTCTCTCCAT 2425  
 Db |||||||  
 QY 2 2 CTCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTCTTAGGATGAACACTCTCTCCAT 61  
 Db |||||||  
 QY 2426 GCGATTGACATATG--ACTTATTTCTAGGGGAAGTCTCTGAGGGGCAACACACAGA 2483  
 Db |||||||  
 QY 62 GCGATTGACATATGAAAGTTATTTGTAGGGGAAGTCTCTGAGGGGCAACACACAGA 121  
 Db |||||||  
 QY 2484 ACCAGGTCCCTCAGCCCAACAGCAGTCTTTTGTGATCCACCCCTCTTACCTTTT 2543  
 Db |||||||  
 QY 122 ACCAGGTCCCTCAGCCCAACAGCAGTCTTTTGTGATCCACCCCTCTTACCTTTT 181  
 Db |||||||  
 QY 2544 ATCAGGATGGGCTGTGTGCTCTTCTGTGGCCATCAGAGACACAGGCAATTAATAT 2603  
 Db |||||||  
 QY 182 ATCAGGATGGGCTGTGTGCTCTTCTGTGGCCATCAGAGACACAGGCAATTAATAT 241  
 Db |||||||  
 QY 2604 TTAACCTTATTTATTAACAAAGTAGAGGGAATCCATTTCTGTGTTGGTGT 2663  
 Db |||||||  
 QY 242 TTAACCTTATTTATTAACAAAGTAGAGGGAATCCATTTCTGTGTTGGTGT 301  
 Db |||||||  
 QY 2664 CTAATATTTGGGTAGGTGGGGATCCCCAACAAATCAGGTGCCCTCAGATAGTGTGTCAT 2723  
 Db |||||||

301	AGTCTGTCAC	TGCCTATAT	TGGTCTCT	CCCGCAG	CGCTGGGT	CTGGTCTGGC	CAATTTACT	TTG	360
1896	CTACACAGG	TAGTATTT	GACAAGAG	CGACATT	TGGCCAA	TAATCTCAG	CGTAGAAA	ACTTCCA	1955
361	CTACACAGG	TAGTATTT	GACAAGAG	CGACATT	TGGCCAA	TAATCTCAG	CGTAGAAA	ACTTCCA	420
1956	GCACATTGG	GGTGGAGG	GGCTGCT	CAC	TGGTCC	CAGCTCC	CGCTCTCT	TGTAGCCCCA	2015
421	GCACATTGG	GGTGGAGG	GGCTGCT	CAC	TGGTCC	CAGCTCC	CGCTCTCT	TGTAGCCCCA	480
2016	TGGGGTGC	CGGGCTG	SCCGCC	AGTTCT	TGCTG	TGCCAA	AGTAATGT	GCTCTCTGCTG	2075
481	TGGGGTGC	CGGGCTG	SCCGCG	CAGG	TCTCT	TGCTGCCA	AGAAATGT	GCTCTCTACTG	540
2076	CCACCTGT	GCTGCT	GAGTGG	TGCGT	AGTC	GCACAG	CTGGGG		2115
541	CCACCTGT	GCTGCT	GAGGAG	CGCTAG	CTGCAC	AACGGGG			580

RESULT	14
BG242597	
LOCUS	linear EST 13-FEB-2001 602334010F1 NCI_CGAP_MamI Mus musculus cDNA clone IMAGE:4482362 S'
DEFINITION	mRNA sequence.
ACCESSION	BG242597
VERSION	BG242597.1 GI:12752412
KEYWORDS	EST.
SOURCE	house mouse. Mus musculus
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgabbs@email.nih.gov
COMMENT	Tissue procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="#">http://image.llnl.gov</a> Plate: LLAMI0319 row: C column: 03 High quality sequence stop: 566.

```

location/xxdml1116
1. 1116
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4482362"
/clone_lib="NCI_CGAP_Maml"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/notes="Organ: Mammary; Vector: pCMV-SPORT6; Site_1: Salt;
Site_2: Notri; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 209 a 349 C 314 G 244 t
ORIGIN
Query Match 15.9%; Score 543.4; DB 12; Length 1116;
Best Local Similarity 84.8%; Pred. No. 1.2e-62;
Matches 644; Conservative 0; Mismatches 111; Indels 4; Gaps 3;

QY 1023 GCTTGGCTTTCCGGAAACCTGGGGCCGCTGTTCCCGGGCTGCACCAAGCTGTGTCGCCGA 1082
|||||
Db 9 GCCTGGCTTTCCGGATCTGGGTACCCCTGTTTCCCGGGCTGCACAGAGCTGTGTCGCCGA 68
|||||
QY 1083 TGCCCGCGCACCCCTGCGCCGCGCTCTTCCTGGCTGAGTGTGCAGCTCGATGGCACTCATGA 1142
|||||

```

Db	69	TGCTCGCACCCCTACGCCGACTCTTTGTGGCTGAGCTGTGCTGAGCTGGATGGCAGCTTATGA	128
QY	1143	CTTTCACGCTGTTTATACAGGATTTCTGTTGGCGAGGGCTGTATACCAGGGCGTGTCCCGAGAG	1202
Db	129	CTTTCACACTGTTCTACAGGACTTCTGTTGGGAGGGGCTGTATACCAGGTTGTATCCAGAG	188
QY	1203	CTGAGCGGGCAGCCGAGGCGCGAGACACTATGATGAAGCGCTTCCGATGGGCGAGCTGG	1262
Db	189	CCGAGCAGGACCCGAGGCGCGAGACACTATGATGAAGCATTTCAATATGGCAGCGCTGG	248
QY	1263	GGCTGTTTCTGTCAGTGGGCAATCTCCCTGGTCTTCTCTGTGTCATGAGCAGCGCTGTGC	1322
Db	249	GGCTCTTCTGTCAGTGGCATCTCCCTGGTCTTCTCCCTGGTCTGTCATGAGCAGCTGGTAC	308
QY	1323	AGCGATTCGGCACTCAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGGCG	1382
Db	309	AGAAGTTGGGCACACGGTCAGTCTATCTGGCCAGTGTGATGACCTTTCTGTGGCTGGCG	368
QY	1383	GTGCCACATCCCTGTCCACAGTGTGGCGTGTGACAGCTTCAGCGCGCTCACCGGGT	1442
Db	369	CTGCCACCTGCCGTGCCACAGCGTGGTGTGACAGCTGAGTGGCTTACCGCTACCGGGT	428
QY	1443	TCACCTTCTAGCCCTGTCAGATCTCTCCCTACACACTGGCTCCCTTACACCGCGAGA	1502
Db	429	TCACCTTCTCGGCTTTCAGATCCTGCCCTTACACGCTGCCCTCCCTTACACCGGTAGA	488
QY	1503	AGCAGTGTCTTCGCCCAATACCGAGGGGACACTGGAGTGTGTCAGTGGAGGACGCC	1562
Db	489	AGCAGTGTCTTCGCCCAATACCGAGGGGACGCTGGAGTGTGTCAGTGGAGGACGCC	548
QY	1563	TGATGACCACTCTCTTCGCCAGGCGCTTAAGCCTGGAGCTCCCTTCCCTTAATGGACAGTGG	1622
Db	549	AGACACCACTCTTCTGCCAGGCGCTTAAGCAGGAGCTCTCTTCCCATATGACAGTGG	608
QY	1623	GTCTGGAGGACAGTGGCTGTCTCCACCTCCACCGCGCTGTGCGGGGCTCTGTGCTGTG	1682
Db	609	GTCTGGGACCAACGAGCGGATCT--GGCCCTCCACTGCACCTCTGTGGGCGCTCTTGTGCG	666
QY	1683	ATGCTCTCGTACGTGT--GTTGGTGGTGGAGCCACCGAGGCGAGTGTTCGGGCGCG	1741
Db	667	ATGTTTCCATGCGAGTGGGTGGGTGGAGCCACCTTTGGGGCGGCTGTGCGGGCGCG	726
QY	1742	GGCATGTGCTGGACCTCGCCATCTCCGATGTGCTTC	1780
Db	727	GGCATTTGCGTGG--CTTGCCATCTTGCGCGGTGCTTC	764
RESULT 15			
BG828841			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			



Db	481	CCACTCCCTCTACTCTCTAGGACTGGGCTGATGAANGCACTGCCCAAAATTTTCCCT	540
Qy	2998	ACCCCAACTTTCCCTACCCCACTTTCCCAACAGCTCCACACCTTCTTTGGAGC	3056
Db	541	ACCCCAACTTTCCCTACCCCACTTTCCCAACAGCTTTCCCAACAGCTTTGGAGC	600
Qy	3057	TACTGC-AGGACCAAGACACAAAGTGGGTTTCCCAAGCTTTTGCATCTCAGCCCC	3114
Db	601	TACTGCCAGGACCCGAAGCCCAAGTGGGTTTCCCAAGCTTTTGGGCTTTTCAGCCC	660
Qy	3115	C-----AGACTATATCTGCTGGGATCTCACAGAACTCAGGAGCACCCTTCG	3169
Db	661	CCCAAGAGATATCTGGGTTTGGGAAACCCCTCCCAAGCTTTTCAGCCCC	720
Qy	3170	CTGAGCTAA 3178	
Db	721	CTGCCCTGA 729	
RESULT 11			
LOCUS	602552833F1 NIH_MGC_15	894 bp mRNA linear EST 21-MAR-2001	
DEFINITION	602552833F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4660496 5', mRNA sequence.		
ACCESSION	BG469520		
VERSION	BG469520.1	GI:13401795	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	NIH-MGC http://mgi.nhl.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC DNA Library Preparation: Ling Hong/Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: NIH Intramural Sequencing Center Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCMI458 row: a column: 09 High quality sequence stop: 671.		
FEATURES	Location/Qualifiers		
source	1..894		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4660496"		
	/clone_lib="NIH_MGC_15"		
	/tissue_type="adenocarcinoma cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"		
BASE COUNT	173 a 272 c 264 g 185 t		
ORIGIN			
Query Match	17.48;	Score 593;	DB 12; Length 894;
Best Local Similarity	97.78;	Pred. No. 3.9e-69;	
Matches 644;	Conservative 0;	Mismatches 10;	Indels 5; Gaps 4;
Qy	1450	CTCAGCCCTGCAGATCCTGCCTACACACTGGCTCCCTCTACACCGGGAGAGCAGGT	1509
Db	2	CTCAGCCCTGCAGATCCTGCCTACACACTGGCTCCCTCTACACCGGGAGAGCAGGT	61
Qy	1510	GTTCCTGCCCAATAACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACACGCTGATGAC	1569
Db	62	GTTCCTGCCCAATAACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACACGCTGATGAC	121
Qy	1570	CAGCTTCTGCCAGGCCCTAAGCCTGGAGCTTCCCTTCCCTAATGGACACGTTGGTCTGG	1629
Db	122	CAGCTTCTGCCAGGCCCTAAGCCTGGAGCTTCCCTTCCCTAATGGACACGTTGGTCTGG	181
Qy	1630	AGGCAAGTGGCTTCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCTGTGATGTCTC	1689
Db	182	AGGCAAGTGGCTTCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCTGTGATGTCTC	241
Qy	1690	CCTACGTGTGTGGTGGTGGAGCCACCGAGGCCAGGTGTTCGGGGCCGGGGCATCT-	1748
Db	242	CCTACGTGTGTGGTGGTGGAGCCACCGAGGCCAGGTGTTCGGGGCCGGGGCATCTC	301
Qy	1749	GCCTGGAGCTGCCCATCTGGATAGTCTTCCCTCTCCAGGTGGCGCCCATCCCTGT	1808
Db	302	GCCTGGAGCTGCCCATCTGGATAGTCTTCCCTCTCCAGGTGGCGCCCATCCCTGT	361
Qy	1809	TTATGGGTCCATTCTCCAGCTCAGCCAGTCTGTCTCTCTCTCTCTCTCTCTCTCTCT	1867
Db	362	TTATGGGTCCATTCTCCAGCTCAGCCAGTCTGTCTCTCTCTCTCTCTCTCTCTCTCT	421
Qy	1868	GGCCTGGGTCTGGTCCCATTTACTTTGTCTACACAGGTAGTATTTGACAAGGACATG	1927
Db	422	GGCCTGGGTCTGGTCCCATTTACTTTGTCTACACAGGTAGTATTTGACAAGGACATG	481
Qy	1928	GCCAATCTACTAGCCTAGAAACTTCCAGCATTTGGGTGGAGGCTGCTCTACTCTGGG	1987
Db	482	GCCAATCTACTAGCCTAGAAACTTCCAGCATTTGGGTGGAGGCTGCTCTACTCTGGG	541
Qy	1988	TCCCAGCTCCCGCTCTCTTTAGCCCCCATGGGGTGGCC--GGGCTGCCGCCCA--GTTTCT	2044
Db	542	TCCCAGCTCCCGCTCTCTTTAGCCCCCATGGGGTGGCC--GGGCTGCCGCCCA--GTTTCT	601
Qy	2045	GTTCGTGCCAAAGTAAATGTGGCTCTCTGCTGCCACCTCTGCTGTGAGGTGCGTAGCT	2103
Db	602	TTGGTGGCAAGTAAATGTGGCTCTCTGCTGACACCTCTGCTGTGAGGTGCGTAGCT	660
RESULT 12			
LOCUS	BI107873	858 bp mRNA linear EST 26-JUN-2001	
DEFINITION	602901816F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5031771 5', mRNA sequence.		
ACCESSION	BI107873		
VERSION	BI107873.1	GI:14558766	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	NIH-MGC http://mgi.nhl.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11088 row: o column: 04 High quality sequence stop: 810.		
FEATURES	Location/Qualifiers		
source	1..858		
	/organism="Mus musculus"		
	/strain="129, C57BL/6J, FVB/N"		
	/db_xref="taxon:10090"		







```

IMAGE:6204253 5', mRNA sequence.
BQ950805
VERSION BQ950805.1 GI:22366283
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 959)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13625 row: d column: 14
High quality sequence start: 2
High quality sequence stop: 490.
FEATURES
source
Location/Qualifiers
1..959
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6204253"
/clone_lib="Lupski_sciatic_nerve"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/Note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dr priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCCGCTCCG-3' and
5'-GACAGTCTAGTACGCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT 139 a 312 c 308 g 200 t
ORIGIN
Query Match 17.7%; Score 604.2; DB 14; Length 959;
Best Local Similarity 97.9%; Pred. No. 1.3e-70;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 1614 GACACGTGGTGTGGAGGAGTGGCTTCTCCACCTCCACCGCGCTCTGGGGGCGCT 1673
Dy 3 GAACCGTGGTGTGGAGGAGTGGCTTCTCCACCTCCACCGCGCTCTGGGGGCGCT 62
Qy 1674 CTGCTCTGATGCTCTCGTACGTGTGCTGGTGGTGGAGCCAGCCAGGCGGTGCTTC 1733
Dy 63 CTGCTCTGATGCTCTCGTACGTGTGCTGGTGGTGGAGCCAGCCAGGCGGTGCTTC 122
Qy 1734 CGGGCCGGGCGATCTGCTGGAGCTCGCCATCTCTGATAGTGGTGGTGGTGGTGGTGG 1793
Dy 123 CGGGCCGGGCGATCTGCTGGAGCTCGCCATCTCTGATAGTGGTGGTGGTGGTGGTGG 182
Qy 1794 TGGCCCGCATCCCTGTTTATGGGTGCTATGTCAGCTCAGCCAGTCTGCTACCTATATA 1853
Dy 183 TGGCCCGCATCCCTGTTTATGGGTGCTATGTCAGCTCAGCCAGTCTGCTACCTATATA 242
Qy 1854 TGGTGTCTGCGGAGCGCTGGTCTGCTGCGCATTTACTTGTACACAGGTAGTATTG 1913
Dy 243 TGGTGTCTGCGGAGCGCTGGTCTGCTGCGCATTTACTTGTACACAGGTAGTATTG 302
Qy 1914 ACAAGAGCGACTTGGCCAAATACTACGCTAGAAAACCTTCCAGCACATTTGGGTGGAGG 1973
Dy 302

```

```

Db 303 ACAAGAGCGACTTGGCCAAATACTACGCTAGAAAACCTTCCAGCACATTTGGGTGGAGG 362
Qy 1974 CTGCTCTACTGGTCCAGCTCCCGCTCCTGTTAGCCCATGGGGCTGCCGGGTGCG 2033
Dy 363 CTTGCTCTACTGGTCCAGCTCCCGCTCCTGTTAGCCCATGGGGCTGCCGGGTGCG 422
Qy 2034 CGCAGTCTTCTGTTGCTGCCAAAGTATATGCTCTCTGCTGCCACCCCTGCTGCTGCTGAG 2093
Dy 423 CGCAGTCTTCTGTTGCTGCCAAAGTATATGCTCTCTGCTGCCACCCCTGCTGCTGCTGAG 482
Qy 2094 GTGCTAGCTGCACAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCT 2153
Dy 483 GTGCTAGCTGCACAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCT 542
Qy 2154 CTGCTAGCTGGAGGCTTCCAAAGGGGGTTCAGTCTGAGCTTATACAGGGAGGCGCAGAA 2213
Dy 543 CTGCTAGCTGGAGGCTTCCAAAGGGGGTTCAGTCTGAGCTTATACAGGGAGGCGCAGAA 602
Qy 2214 GGCTCCATGCACCTGGAATGCGGGG 2238
Dy 603 GGCTCCATGCATGGAATGCGGGG 627

```

```

RESULT 9
BF972601 850 bp mRNA linear EST 22-JAN-2001
LOCUS 602243025F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4331407 5',
DEFINITION mRNA sequence.
ACCESSION BF972601
VERSION BF972601.1 GI:12339816
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 850)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Incyte Genomics, Inc.
Found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCML196 row: a column: 08
High quality sequence stop: 675.
FEATURES
source
Location/Qualifiers
1..850
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4331407"
/clone_lib="NIH_MGC_46"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
III RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 128 a 266 c 261 g 195 t
ORIGIN
Query Match 17.4%; Score 594.6; DB 12; Length 850;
Best Local Similarity 92.9%; Pred. No. 2.4e-69;
Matches 733; Conservative 0; Mismatches 44; Indels 12; Gaps 10;

```



QY	2815	GATAATTCCAAATGCTGTACCCAAAGGTTAGGGTGTGAAGGAAGGTAGAGGGTGGGGCT	2874
Db	541	GATAATTCCAAATGCTGTACCCAAAGGTTAGGGTGTGAAGGAAGGTAGAGGGTGGGGCT	600
QY	2875	TCAGGTCTCAACGGC--TTCCCTAACACACCCCTCTCTCTTGTCGCCAG-CCTGGTTCCCC	2932
Db	601	TCAGGTCTCAACGGCTTTCCTAACCAACCCCTCTCTCTTGTCGCCAGCCCTGTTCCCC	660
QY	2933	CACCTTCCACTCCCGCTACTCTCTAGACGTGGGCTGATGAAGGACACTGCCCAAAATTT	2992
Db	661	CACCTTCCACTCCCGCTACTCTCTAGACGTGGGCTGATGAAGGACACTGCCCAAAATTT	720
QY	2993	CCCTTACCCGCCAAC--TTTCCGCTACCCCAACTTTTCC	3030
Db	721	CCCTTACCCGCCAACITTTTCCCTTACCCCAACTTTTCC	760
RESULT 6			
LOCUS	BG469487		
DEFINITION	602532933F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4660703 5',	800 bp	linear EST 21-MAR-2001
ACCESSION	BG469487		
VERSION	BG469487		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Authors	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
	1. (bases 1 to 800)		

```

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
High quality sequence stop: 741.
Location/Qualifiers
1. 800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4660703"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
135 a 245 c 239 g 181 t

BASE COUNT
ORIGIN

Query Match 19.6%; Score 667.4; DB 12; Length 800;
Best Local Similarity 97.0%; Pred. No. 6.6e-79;
Matches 744; Conservative 0; Mismatches 16; Indels 7; Gaps 6;

1558 CAGCCTGATGACCAGCTTCCTGCCAGGCCCTTAAGCCTGGAGCTCCCTTCCTAATGCACA 1617
|||||
2 CAGCCTGATGACCAGCTTCCTGCCAGGCCCTTAAGCCTGGAGCTCCCTTCCTAATGCACA 61
|||||

1618 CGTGCGTCGAGCCACTGGCTGCGCCCACTGCACGCCCCCGCCGCTCCGCCCCCGAGCC 1677

```

1077

QY 1022 CGCTGGCTTCCGGA--CCTGGCGCCCTGCTCCCGGCTGCACAGCTGTGCTGCC 1079  
 |||||  
 Db 660 CGCTGGCTTCCGGAACCTGGGGCGCCCTGCTTCCCGGCTGCACAGCTGTGCTGCC 719  
 |||||  
 QY 1080 GCATGCCCGCCACCCCTGCGCGGCTCTTGCTGGCTGAGCTGTGACAGTGG--ATGSCACT 1137  
 |||||  
 Db 720 GGCATGCCCGGAACCTGCGCGGCTCTTGCTGGCTGAGCTGTGACAGTGGCAACT 779  
 |||||  
 QY 1138 CATGACCTTCACGCTGTTTACAGG--ATTTCGTGGCGGAGGGCTGT-ACCAGGGCT 1194  
 |||||  
 Db 780 CATGACCTTCAGCTGCTTTCACCGGCATCTCGCTGGCGAGCGGCTGTGCCACAGGGCT 839  
 |||||  
 QY 1195 GCCCAGAGCTGAGCGGGGACC--GAGCGCGGAGACACTATGATGAAGCGTTCGGATGG 1253  
 |||||  
 Db 840 TGCAGAGCTGAGCGGGGACCGAAGCGGAGAACCTATGCTGAACGCGTCCGCAT-G 898  
 |||||  
 QY 1254 GCAGCTGGGGCTGTTCTCCTGCAGTGC 1279  
 |||||  
 Db 899 GCCCGCTGGGCTGTCTCTGACAGCG 924  
 |||||

## RESULT 4

BM915082 1025 bp mRNA linear EST 12-MAR-2002  
 LOCUS AGENCOURT\_6702317 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5481218  
 5', mRNA sequence.  
 ACCESSION BM915082  
 VERSION BM915082.1 GI:19365461  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

## REFERENCE

1 (bases 1 to 1025)

## AUTHORS

NIH-MGC <http://mgi.nci.nih.gov/>.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: L1CM2005 row: f column: 03

High quality sequence start: 35

High quality sequence stop: 657.

## FEATURES

## source

1..1025  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5481218"  
 /clone.lib="NIH\_MGC\_41"  
 /tissue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pOPB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."  
 BASE COUNT 237 a 273 c 252 g 263 t  
 ORIGIN

## Query Match

Best Local Similarity 22.0%; Score 748.6; DB 14; Length 1025;

Matches 802; Conservative 0; Mismatches 24; Indels 9; Gaps 3;

QY 1799 CCATCCCTGTTTGGGCTCCATTGTCAGCTGCACGCTGTGCTACGTGCTATATGTT 1858

Db 12 CCATCCCTGTTTGGGCTCCATTGTCAGCTGCACGCTGTGCTACGTATATGTTG 71  
 |||||  
 QY 1859 TCTGCCGAGCGCTGGGCTGTGTCGCTTACTTTTCTACACAGTAGTATTTGACAAG 1918  
 |||||  
 Db 72 TCTGCCGAGCGCTGGGCTGTGTCGCTTACTTTTGTACACAGTAGTATTTGACAAG 131  
 |||||  
 QY 1919 AGCGACTTTGGCCAAATACTACAGCGCTAGAAAACCTTCCACACATTTGGGCTGAGGCGCTGC 1978  
 |||||  
 Db 132 AGCGACTTTGGCCAAATACTACAGCGCTAGAAAACCTTCCACACATTTGGGCTGAGGCGCTGC 191  
 |||||  
 QY 1979 CTCACCTGGTCCAGACTCCCGCTCCTTTAGCCCCATGGGCTGCCGGGCTGCCGCCA 2038  
 |||||  
 Db 192 CTCACCTGGTCCAGACTCCCGCTCCTTTAGCCCCATGGGCTGCCGGGCTGCCGCCA 251  
 |||||  
 QY 2039 GTTTCCTGCTGCCAAAAGTAATGTGCTCTGTGCTGCACCCCTGTGCTGCTGAGGTGCG 2098  
 |||||  
 Db 252 GTTTCCTGCTGCCAAAAGTAATGTGCTCTGTGCTGCACCCCTGTGCTGCTGAGGTGCG 311  
 |||||  
 QY 2099 TAGCTGCACAGCTGGGGCTGGGGCTCCCTCTCTCTCCCAAGTCTTAGGGCTGCC 2158  
 |||||  
 Db 312 TAACTGCACAGCTGGGGCTGGGGCTCCCTCTCTCTCCCAAGTCTTAGGGCTGCC 371  
 |||||  
 QY 2159 TCACCTGGAGCGCTTCCAGGGGCTTTCAGCTCTGGACTTATACAGGAGGCGCAGAGGCT 2218  
 |||||  
 Db 372 TCACCTGGAGCGCTTCCAGGGGCTTTCAGCTCTGGACTTATACAGGAGGCGCAGAGGCT 431  
 |||||  
 QY 2219 CCATGCACCTGGAATGCGGGGCTGCGAGGTGATTACCCAGGCTCAGGGTTAACAGCTA 2278  
 |||||  
 Db 432 CCATGCACCTGGAATGCGGGGCTGCGAGGTGATTACCCAGGCTCAGGGTTAACAGCTA 491  
 |||||  
 QY 2279 GCCTCTAGTTGAGACACACCTAGAGAGGGGTTTTGGAGCTGAATAAAGTCACTCACC 2338  
 |||||  
 Db 492 GCCTCTAGTTGAGACACACCTAGAGAGGGGTTTTGGAGCTGAATAAAGTCACTCACC 551  
 |||||  
 QY 2339 TGGTTTCCCATCTCTAAGCCCTTAACTGACGCTTCTGTTTAAATGTAGCTCTTGCATGG 2398  
 |||||  
 Db 552 TGGTTTCCCATCTCTAAGCCCTTAACTGACGCTTCTGTTTAAATGTAGCTCTTGCATGG 611  
 |||||  
 QY 2399 AGTTCTAGGATGAACACCTCTCCATGGGATTGAACATATGAAGTTATTGTAGGG 2456  
 |||||  
 Db 612 AGTTCTAGGATGAACACCTCTCCATGGGATTGAACATATGAAGTTATTGTAGGG 671  
 |||||  
 QY 2457 AGAGTCTGAGGGGCAACACACAGAGCCAGTCCCTCAGCCACACACACTGCTTTT 2516  
 |||||  
 Db 672 AGAGTCTGAGGGGCAACACACAGAGCCAGTCCCTCAGCCACACACACTGCTTTT 731  
 |||||  
 QY 2517 TGCTGATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTTGGTCC---TTCTGTT 2573  
 |||||  
 Db 732 TGCTGATCCACCCCTCTTACCTTTTATCAGGATGTGGCTGTTGGCTTCTGTT 791  
 |||||  
 QY 2574 GCATCAGACAGACAGGCACTTT---AAATATTTAACTTATTTTAAACAA 2624  
 |||||  
 Db 792 GCCTCCACAAAACAGGGGTTATTAATAAAATTTAACTTATTTTAAACAA 846  
 |||||

## RESULT 5

## BM950912

## LOCUS

## DEFINITION

BM950912

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

## REFERENCE

1 (bases 1 to 916)

## AUTHORS

NIH-MGC <http://mgi.nci.nih.gov/>.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2003, 15:13:08 ; Search time 4204 Seconds  
(without alignments)  
13136.682 Million cell updates/sec

Title: US-09-759-143-110

Perfect score: 3410

Sequence: 1 gggaccagcctgcgcgcg.....aaaaaaaaaaaaaaaaaaaa 3410

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estin:\*\*

4: em\_estmu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_htc:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_htc:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: gb\_gss:\*\*

18: em\_gss\_hum:\*\*

19: em\_gss\_inv:\*\*

20: em\_gss\_pln:\*\*

21: em\_gss\_vit:\*\*

22: em\_gss\_fun:\*\*

23: em\_gss\_mam:\*\*

24: em\_gss\_mus:\*\*

25: em\_gss\_othr:\*\*

26: em\_gss\_pro:\*\*

27: em\_gss\_rod:\*\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	784.2	23.0	885	14	BQ948390
2	787.2	22.5	1063	14	BM915527
3	749.6	22.0	946	12	AG469889
4	748.6	20.6	1025	14	BM915082
5	703.2	20.6	916	14	BQ950912
6	667.4	19.6	800	12	AG469487

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	660.2	19.4	718	12	BE867241
8	604.2	17.7	959	14	BQ950805
9	594.6	17.4	850	12	BF972601
10	593.2	17.4	875	14	BQ942028
11	593	17.4	894	12	AG469520
12	588	17.2	858	13	BI107873
13	564	16.5	1060	14	BM914562
14	543.4	15.9	1116	12	AG424297
15	538.2	15.8	1035	12	AG424297
16	534.8	15.7	589	12	AG469586
17	525	15.4	715	12	AG424297
18	512.8	15.0	1667	14	BM912193
19	508.6	14.9	872	12	AG464609
20	499.6	14.7	786	12	AG424297
21	463	13.6	901	13	BI650119
22	462.6	13.6	482	9	AF109299
23	459	13.5	969	12	AG424297
24	456.8	13.4	491	9	AA631143
25	456.2	13.4	929	12	BF785813
26	455.2	13.3	592	10	AA412402
27	453	13.3	934	14	BQ934815
28	449.2	13.2	469	9	AI703348
29	438.6	12.9	549	12	AG424297
30	435.4	12.8	448	10	BE74096
31	433.8	12.7	700	12	BF581244
32	419.6	12.3	700	10	BB610495
33	419.2	12.3	759	12	BF789072
34	409.4	12.0	692	10	BB627667
35	400.2	11.7	844	12	AG424297
36	398.6	11.7	412	10	AW175665
37	395.8	11.6	418	12	BF922235
38	391	11.5	430	9	AF109303
39	389.4	11.4	630	13	BI145201
40	381.2	11.2	650	10	BB627844
41	369.4	10.8	537	10	AW787124
42	366	10.7	388	9	AA579486
43	359.2	10.5	906	12	BF680993
44	353	10.4	370	10	AW135465
45	352.8	10.3	509	10	BB701488

# ALIGNMENTS

RESULT 1  
BQ948390  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BQ948390 885 bp mRNA linear EST 21-AUG-2002  
AGENCOURT\_8803099 Lupski\_sciatic\_nerve Homo sapiens cDNA clone  
IMAGE:6198823 5', mRNA sequence.  
BQ948390  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 885)  
NIH-MGC <http://mgc.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999).  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM13611 row: b column: 08  
High quality sequence stop: 615.  
Location/Qualifiers  
1. .885



Db 243 erProGlyGlyLysGlyGlu-----MetGlyProAlaG 254  
QY 2025 GGCAGCCCATGGGCTAACAGGAGGGGAGGCTGGACCCAGTGGAGGAGCCCTCCAC 1966  
Db 254 lytleProGlyAlaProGlyLeuMetGlyAlaArgGlyPro-----ProGlyProAlaG 272  
QY 1965 CCCAATGTGTGGAAGTTTCTACGTG-----AGTATTGGCCAAAGTCGCTTGTCAAA 1910  
Db 272 lyAlaAsnGlyAlaProGlyLeuArgGlyGlyAlaGlyGluProGlyLysAsn-----289  
QY 1909 TACTACCTGTGTAGCAAGTAATGGCGACACACCA-----G 1871  
Db 290 -----GlyAlaGlyGlyGluProGlyProArgGlyGluArgGlyGluAlaG 305  
QY 1870 GCCTGGCGACACACCATATAGGCAGTGCAGACTGGCTGAGCTGGACAAATGAGCCCAT. 1811  
Db 305 lytleProGlyValProGlyAlaLysGlyGlyAlaAspGlyLysAspGlySerProGlyGluP 325  
QY 1810 AACAGGATGGGCCACCTGGGACACGAG-----AGGCATATCCAGGATGGCGA-----1758  
Db 325 roGlyAlaAsnGlyLeuProGlyAlaAlaGlyGluArgGlyAlaProGlyPheArgGlyP 345  
QY 1757 -----GTTCCAGGCAGATGCCCGCCCGGAACACCCCTCGCTGGTGGCTCACCCAC 1703  
Db 345 roAlaGlyProAsnGlyleProGlyGlyLysGlyProAlaGlyGluArgGlyAlaProG 365  
QY 1702 CACCACAGTACGAGACATCACAGGAGGAGGCCCGCAGAGC-----GCGGGTGGAG 1650  
Db 365 lyProAlaGlyProArgGlyAlaAlaGlyGluProGlyArgAspGlyValProGlyGlyP 385  
QY 1649 GTGGAGAGCCCACTGCTCCAGCACCCAGCTGTCCATTAGGAGGGAGTCCAGGCT 1590  
Db 385 roGlyMetArgGlyMetProGlySerProGlyGlyProGlySerAspGlyLysProGly- 404  
QY 1589 TAGGCTGGCAGGAGCTGTCTCATAGGCTGTCTCTCCTCAGTGCATGACCTCCAGTGTCC 1530  
Db 405 --ProProGly- SerGlnGlyGlySerGlyArgPro-----415  
QY 1529 CTCGGTATTGGGAGGAACACCTGCTTCTCCGGTGTGTAGAGGAGGCCAGTGTGTAGG 1470  
Db 416 -----GlyProGlyProSerGlyProArgGlyGlnPro-----427  
QY 1469 GCAGGATGTCAGGCTGAGNAGGTGAACCCGGTGGGCGGTGAGCTGTCCACACGG 1410  
Db 428 --GlyValMetGlyPhe-----ProGly 434  
QY 1409 CCACACTGTGGACAGCATGTGACCCGCGCAGCACAGGAAAGCTGCCACACTGSCCA. 1350  
Db 435 ProLysGlyAsnAspGly-----AlaProGlyLys-----444  
QY 1349 AATAGACTGTCTGAGTCCGCAATCGCTGCACCCGGTCCATGACACAGAGAAGACCA 1290  
Db 445 -----AsnGlyGluArgGlyGlyProGlyGlyProGlyProGlnGlyProPro 460  
QY 1289 GGGAGATGGCCACTGCAGGA-----ACAGCCCGCAGGTCGCCCA 1251  
Db 461 GlyLysAsnGlyGluThrGlyProGlnGlyProProGlyProThrGlyProGlyGlyAsp 480  
QY 1250 TCCGAACCCCTTCATCATAGTGTCTCCGGGCTCGGTCCCGGCTCAGCTCGGCGACGC 1191  
Db 481 Lys-----GlyAspThrGlyProGlyProGlnGlyLeuGlnGlyLeu 495  
QY 1190 CTGTGTACA-----1182  
Db 496 ProGlyThrGlyGlyProProGlyGluAsnGlyLysProGlyGluProGlyProLysGly 515  
QY 1181 -----GCCCTCGCCACCA 1167  
Db 516 AspAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAlaProGlyGluArg 535  
QY 1166 AATCCGTATAAACAGCTGAAGTGCATGAGTGCATCCAGCTGCACAGCTCAGCCACGA 1107  
Db 536 GlyPro-----ProGlyLeuAlaGlyAlaProGlyLeu 546

QY 1106 AGAGCCGGCGCAGGTGGCGGCATGCGCAGCACAGCTGGTGCAGCCGGGGAACAGGG 1047  
Db 547 ArgGlyGlyAlaGlyProProGlyGlyGlyLysGlyAlaAlaGlyProProGly 566  
QY 1046 CCCCAGGTTC-----1035  
Db 567 ProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGluArgGlyGlyLeu 586  
QY 1034 -----GGAAGCCCAAGCGCGGCATGACAGCAGT 1002  
Db 587 GlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGlyAlaAspGlyVal 606  
QY 1001 GGGGCGACA-----AGGAGGGCGCAGACAGCCCTTCTGCTGGCTCGTGGGCCCCAGCG 948  
Db 607 ProGlyLysAspGlyProArgGlyProThr-----GlyProLys 619  
QY 947 CTGCTCTCTCAGCCACAGCAGTGTGGCTGCTACGAGGTGAGGAGATGAGGTGAGCA 888  
Db 620 GlyProProGlyProAla-----GlyLysProGlyAspLys 631  
QY 887 GGCACAAAGAGCAGCTCTCTGGTGGTCCAGGT---AGGGGCGCAGGCGCTGTGTCC 831  
Db 632 GlyGluGlyGlyAla---ProGlyLeuProGlyIleAlaGlyPro-----645  
QY 830 AGTCAATGCGAGGAGGTAGTCCAGCAGCAGCCCAAGACTGATCATGAAGGCATAGA 771  
Db 646 -----ArgGlySerProGlyGlu-----ArgGlyGlu 654  
QY 770 CAGAGTAGCTGGCCACAGTGT---CCGGT---CCGGAAGAGGTGACAGAGCAGGG 717  
Db 655 ThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsnGlyGluProGly 674  
QY 716 CTTCCAGTGGTGAACACACCTGGCCACAGTCCAGCAGCAGCCCGCAGGATGA 657  
Db 675 GlyLysGlyGluArgGlyAlaProGlyGlyLysGlyGlyGlyPro---ProGlyVal 693  
QY 656 GCAGTCCAGCTCCAGG-----GCCTGGATCCGGGCGCAGCAGCAGCCCTGTAGCC 606  
Db 694 AlaGlyProProGlyGlySerGlyProAlaGlyProProGlyProGlnGlyValLysGly 713  
QY 605 AGCCGCGCTTGGATGAGAAAGCTCAGCAGGATGCCAAGGACAGCAGTCCCGCATGA 546  
Db 714 -----GluArgGlySerProGlyGlyPro-----721  
QY 545 AGGCGCGCGCGCCATAGCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486  
Db 722 -----GlyAlaAlaGlyPhe-----ProGlyAlaArg 730  
QY 485 CACAGACCCAGCCCGCAGCTGGACCAATGCCAGCAGCAGTGTGATGATGATGATGATGAT 426  
Db 731 GlyLeuProGlyPro-----ProGlySerAsnGlyAsnProGly 743  
QY 425 CCCCCTTCCAGCAGCAGGCGGCACATAGGTGATGCTGGGCGCAACACACCTCCA 366  
Db 744 ProProGlyProSerGlySer-----Pro 751  
QY 365 GGCACAAAGTTAGCAGGTTGACCAAGAGCTGGGCTTTCCGGTCCCGCAGCAGGCGGC 306  
Db 752 GlyLys-----AspGlyProProGlyProAlaGlyAsnThrGlyAlaPro-----Gly 767  
QY 305 TCACCCACAGCTCTTGACCATAGTGG-----CCAGCGGGTAGGCTGAGGCTCAGG 258  
Db 768 SerProGlyValSerGlyPro-LysGlyAspAlaGlyGlnProGlyGlyLysGlySerPr 787  
QY 257 GGGCCCTTCCAGCAGTCCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 212  
Db 787 oGlyAlaGlnGlyPro-----GlyAlaProGlyProLeuGly11 801  
QY 211 -----GAGCTGGCGCTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 171  
Db 801 eAlaGlyIleThrGlyAlaArgGlyLeuAlaGlyProProGlyMetProGly---ProAr 820

QY 170 GAATCAGCCAGCGCCATTTCTGCCAGCCCTTGTGGCGGTCCAGCTTCTCAG-----116  
 Db 820 gglySerProGlyProGlnGlyValLysGlyGluSerGlyLysProGlyAlaAsnGlyLe 840  
 QY 115 -----CCATGCTCAACACCTGCTGCTGTGGGCGCACCTCA 81  
 Db 840 userGlyGluArgGlyProGlyProGlnGlyLeuProGlyLeuAlaGly---Thra 859  
 QY 80 GTGGGGACAGCTCTCACTCACTCAGATCTCTGC-----49  
 Db 859 laglyGluProGlyArgAspGlyAsnProGlySerAspGlyLeuProGlyArgAspGlyS 879  
 QY 48 -----CGAGGCGCGCGGCTGTCAACCCCGAGCC 22  
 Db 879 erProGlyGlyLysGlyAspArgGlyGluAsnGlySerProGlyAla 894

## RESULT 43

US-09-219-849-49  
 : Sequence 49, Application US/09219849  
 : Patent No. 6150081  
 : GENERAL INFORMATION:  
 : APPLICANT: VAN HEERDE, GEORGE V.  
 : APPLICANT: VAN RIJN, ALEXIS C.  
 : APPLICANT: BOWMSTRA, JAN B.  
 : APPLICANT: DE WOLF, FREDERIK A.  
 : APPLICANT: MOOBROEK, ANDREAS  
 : APPLICANT: WERTEN, MARC W.T.  
 : APPLICANT: WIND, RICHELIE D.  
 : APPLICANT: VAN DEN BOSCH, TANJA J.  
 : TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN  
 : TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE  
 : FILE REFERENCE: 2728-2  
 : CURRENT APPLICATION NUMBER: US/09/219,849  
 : CURRENT FILING DATE: 1998-12-23  
 : NUMBER OF SEQ ID NOS: 50  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 49  
 : LENGTH: 822  
 : TYPE: PRT  
 : ORGANISM: Artificial Sequence  
 : FEATURE:  
 : OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 : OTHER INFORMATION: amino acid sequence  
 US-09-219-849-49

## Alignment Scores:

Pred. No.: 6.84e-19 Length: 822  
 Score: 351.00 Matches: 244  
 Percent Similarity: 32.94% Conservative: 37  
 Best Local Similarity: 28.60% Mismatches: 328  
 Query Match: 5.64% Indels: 245  
 DB: 4 Gaps: 43

US-09-759-143-110 (1-3410) x US-09-219-849-49 (1-822)

QY 2131 AGAGGAGCGCCAGCCCGCCAGCTGTGCAGTACGACCTCAGCAGCAGCGGTGCGCAGC 2072  
 Db 8 ArgGlyLeuProGlyProGlyAlaPro---GlyProGlnGlyPheGlnGlyProPro 26  
 QY 2071 AGAGAGCCCATTTACTTTTGGCAGCAACAGAAACTGGCGCCAGCCCGCCAGCCCATGGG 2012  
 Db 27 GlyGluProGlyGluProGlyGlySerGlyProMetGlyProArgGlyProGlyPro 46  
 QY 2011 GCTAACAGGAGCGGGGAGCTGGGA-----CCCAAGTACGAGCGC 1973  
 Db 47 ProGlyLysAsnGlyAspGlyGluAlaGlyLysProGlyArgProGlyGlyArgGly 66  
 QY 1972 CTTCCA-----CCCCAATGTGCTGGGAAGTTTCTACGCTCA 1937  
 Db 67 ProProGlyProGlnGlyAlaArgGlyLeuProGlyThraGly-----81  
 QY 1936 GTATTTGGCCAAAGTCGCTCTTGTCAANAATACCTGTGTAGTACAAAGTAATGGCGCAGC 1877

Db 82 LeuProGlyMetLysGlyHisArgGlyPheSerGlyLeuAspGlyAlaLysGlyAspAla 101  
 QY 1876 ACCCAGGCTCGGGCAGACACCATATAGSCAGTGA---CAGACTGGCTGAGCTGGACAA 1821  
 Db 102 GlyProAlaGlyProLysGlyGluProGlySerProGlyGluAsnGlyAlaProGlyGln 121  
 QY 1820 TGGAGCCCATAAACAGGATG-----GGCCACCTGGGACAGCA 1782  
 Db 122 MetGlyPro---ArgGlyLeuProGlyGluArgGlyArgProGlyProGlyThraAla 140  
 QY 1781 GGAAGGCACTATCCAGGATGGGAGGTCCAGCAGATCCCGGCGCCGGAACACCTGTGG 1722  
 Db 141 GlyAlaArgGly-AsnAspGlyAlaValGlyAlaGlyPro-ArgGlyProThr---G 159  
 QY 1721 CCTCGTGGCTCACCCACACCATCGGAGACATCACAGCAGCAGCC---CCGC 1665  
 Db 159 lyProThrGlyProProGlyPheProGlyAlaValGlyAlaLysGlyGluAlaGlyProG 179  
 QY 1664 AGAGCGCGGTGGAGGT-----GGGAGCAGG-----CCACTGCTCCAGCAC 1623  
 Db 179 lnglyAlaArgGlySerGlyGluProGlnGlyValArgGlyGluProGlyProGlyP 199  
 QY 1622 CCAGCTGCTCATTAGG---AAGGAGCTCCAGGCTTAGG-----CCTGGCAGGAGC 1572  
 Db 199 roAlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGlnProGlyAlaLys- 218  
 QY 1571 TGGTCATCAGGCTGTCTCTCACTGTAGCACCTCCAGTCTCCCTCGGTATTGGGCAGGA 1512  
 Db 219 GlyAlaAsnGlyAlaPro-----Glylle---AlaGly 228  
 QY 1511 ACACCTGCTTCTCCGGTGTAGAGGAGGCCAGTGTGTAGGGCAGGATTCAGGGGTG 1452  
 Db 229 AlaProGlyPheProGlyAlaArgGly-----ProSerGlyPro-GlnGlyPr 244  
 QY 1451 AGAGGTCAACCCGCTGAGGCGGTGAAGCTGTCCACCAGCCACACTGTGGGCAGCAGC 1392  
 Db 244 oSerGlyProGlyProLysGlyAsnSerGlyGluProGly---AlaProGlyAsnLy 263  
 QY 1391 ATGTGGCACCGGACCCACAGGAAAGCTGCCACTGTGCCAAATAGACTGCTCGAGTGC 1332  
 Db 263 sGlyAspThrGlyAlaLysGlyGluProGlyAlaThrGly-----276  
 QY 1331 CGAATCGCTGCACACCGGCTCCATGACCAGAGAGAACAGCAGGAGATGCCGCACTGCA 1272  
 Db 277 -----ValGlnGlyPro---ProGlyProAlaGlyGluGlyLysArgGlyAlaArg 293  
 QY 1271 GCAACA-----GCCCGAGCTGCCCATCCGAAACGC-----1242  
 Db 294 GlyGluProGlyProSerGlyLeuProGlyProGlyGluArgGlyGlyProGlySer 313  
 QY 1241 -----CTTCATCATAGTGTCTCCGGGCTCGGTGCGCGGTCACTGCTGGGCA 1194  
 Db 314 ArgGlyPheProGlyAlaAspGlyValAlaGlyProLysGlyProSerGlyGluArgGly 333  
 QY 1193 CGCCCTGTACAGCCCTCGCCCAAGAAATCCGTGTAAACAGCGGTCAAGTGCATGAGTG 1134  
 Db 334 AlaProGlyProAlaGlyProLysGlySerPro-----344  
 QY 1133 CCATCCAGCTGCACAGCTCAGCCCAAGAGCGCGGCGGATGCGGCGATGCGGCGAGC 1074  
 Db 345 -----Gly 345  
 QY 1073 ACAGCTGTGTGAGCCGGGGAAGAGAGCGGCCAGGT-----TCCGGAAGACCA 1026  
 Db 346 GluAlaGlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGlySerPro 365  
 QY 1025 AGCGGGCCCGCATGACAGCAGTGGCGGACAGAGAGGGGGCCGACAGCCCTTCTCTG 966  
 Db 366 GlySerProGlyProAsp-----GlyLysThrGlyProGlyProAlaGlyGlnAsp 383  
 QY 965 GCTCGGTGGGCGCCAGGCGCTGCT-----942

384	GlyArgProGlyProAlaGlyProProGlyAlaArgGlyGlnAlaGlyValMetGlyPhe	403
941	CCTCAGCCACCAGCAGTGTGCTCCTACGCAGGTGAGGAAGATGAGGCTGAGCAGGCCAA	882
404	ProGlyProLysGlyThrAla-----GlyGluProGlyLysAlaGlyGlu	418
881	AGAGGCATCTCT-----	870
419	ArgGlyLeuProGlyProProGlyAlaValGlyProAlaGlyLysAspGlyGluAlaGly	438
869	---CCTGGTCCAGCT---AGGGGCCAGGGCAGTGTGTCCTCAGTCAATGGCAGCA	816
439	AlaGlnGlyAlaProGlyProAlaGlyGluArgGlyGlnGlyProAla	458
815	GGA-----GGTAGCCAGGCAGGCCCAAGCACTGATCATGA	780
459	GlySerProGlyPheGlnGlyLeuProGlyProAlaGlyProGlyGluAlaGlyLys	478
779	AGGCATACAGATAGGCTCGGCGCACAGTGGT---CCGGGTCC-----GGAAGAGT	729
479	ProGlyGluGlnGlyValProGlyAspLeuGlyAlaProGlyProSerGlyAlaArgGly	498
728	CAGAGCAGGSCCTCCAGTGCAGTGAAGCACACCTGGCCACAGAGTCCAGCAGCCCA	669
499	GluArg---GlyPheProGlyGluArgGlyValGlnGlyProProGlyProAlaGlyPro	517
668	CGCCAGGATGAGCAGTGCCA-----GCTCCAGGG	639
518	ArgGlyAsnAsnGlyAlaProGlyAsnAspGlyAlaLysGlyAspThrGlyAlaProGly	537
638	GCCTGGGATCCGGGCACAGCCCTGCTAGCCAGCGGCCCTTGGATGAGAAAGGC	579
538	Ala-----ProGlySerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArgGly	555
578	TCAGCAGATGCCNAGGACACTGCCAGATCAAGGCGCGCGGCCCATACGCTCCAC	519
556	AlaAlaGlyLeuProGly-----ProLysGlyAspArgGlyAspAlaGlyProLysGly	573
518	GCCAGTGTCTACTGGCTCAGCTAGGACGGGACACAGACAGCCAGCCAGCAGCTGGACCA	459
574	AlaAspGlySerProGlyLysAspGlyAlaArgGlyLeuThrGlyProLys-----	590
458	TGCCAGCACCATGTCATGAACCTCTCCTACTACCCCACTTCCAGCAGAGGGGCA	399
591	-----GlyProGlyProAlaGlyAlaProGly	600
398	CATAGTGATGCTCGCGCCAAACACACTCCAGC-----	363
601	AspLysGlyGluAlaGlyProSerGlyProGlyProThrGlyAlaArgGlyAlaPro	620
362	-----CAAAGTTAGCAGGTGACCAAGAGCTGGCTTTCCGGT---GCCGCAGCA	312
621	GlyAspArgGlyGluAlaGlyProProGlyProAlaGlyPheAlaGlyProProGlyAla	640
311	GCGGCTCACCCACAGCTCT-----	291
641	AspGlyGlnProGlyAlaLysGlyGluProGlyAspThrGlyValLysGlyAspAlaGly	660
290	-----GGACCATAGTGGCCAGCGGG---TAGGGCTAGGGGGCGTTCAGGCATCC	240
661	ProProGlyPro-AlaGlyProAlaGlyProProGlyProLysGlyAsnValGlyAlaPr	680
239	AGAACTGCTCTCGCTCTGCTGCCAGAGCTGCGGCCCTCTCCTCTTGTCTGCCGCCA	180
680	oGly-----ProLysGlyProArgGlyAlaAlaGlyProPro-GlyAlaThr-	695
179	ACTGCCTAGGAATCAGCCAGCGGCCCAATTTCTGCCAGCCCTTGGTCCGCGTCCAGCTTC	120
696	-----GlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSerGlyAsnA	713
119	TCAGCCCATGCTCAACACTGCTGCTGGGGCACCTCAGT-----GGGG	75
713	lacyProProGlyProProGlyProValGlyLysGluGlyProArgGlyG	733
384	GlyArgProGlyProAlaGlyProProGlyAlaArgGlyGlnAlaGlyValMetGlyPhe	403
941	CCTCAGCCACCAGCAGTGTGCTCCTACGCAGGTGAGGAAGATGAGGCTGAGCAGGCCAA	882
404	ProGlyProLysGlyThrAla-----GlyGluProGlyLysAlaGlyGlu	418
881	AGAGGCATCTCT-----	870
419	ArgGlyLeuProGlyProProGlyAlaValGlyProAlaGlyLysAspGlyGluAlaGly	438
869	---CCTGGTCCAGCT---AGGGGCCAGGGCAGTGTGTCCTCAGTCAATGGCAGCA	816
439	AlaGlnGlyAlaProGlyProAlaGlyGluArgGlyGlnGlyProAla	458
815	GGA-----GGTAGCCAGGCAGGCCCAAGCACTGATCATGA	780
459	GlySerProGlyPheGlnGlyLeuProGlyProAlaGlyProGlyGluAlaGlyLys	478
779	AGGCATACAGATAGGCTCGGCGCACAGTGGT---CCGGGTCC-----GGAAGAGT	729
479	ProGlyGluGlnGlyValProGlyAspLeuGlyAlaProGlyProSerGlyAlaArgGly	498
728	CAGAGCAGGSCCTCCAGTGCAGTGAAGCACACCTGGCCACAGAGTCCAGCAGCCCA	669
499	GluArg---GlyPheProGlyGluArgGlyValGlnGlyProProGlyProAlaGlyPro	517
668	CGCCAGGATGAGCAGTGCCA-----GCTCCAGGG	639
518	ArgGlyAsnAsnGlyAlaProGlyAsnAspGlyAlaLysGlyAspThrGlyAlaProGly	537
638	GCCTGGGATCCGGGCACAGCCCTGCTAGCCAGCGGCCCTTGGATGAGAAAGGC	579
538	Ala-----ProGlySerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArgGly	555
578	TCAGCAGATGCCNAGGACACTGCCAGATCAAGGCGCGCGGCCCATACGCTCCAC	519
556	AlaAlaGlyLeuProGly-----ProLysGlyAspArgGlyAspAlaGlyProLysGly	573
518	GCCAGTGTCTACTGGCTCAGCTAGGACGGGACACAGACAGCCAGCCAGCAGCTGGACCA	459
574	AlaAspGlySerProGlyLysAspGlyAlaArgGlyLeuThrGlyProLys-----	590
458	TGCCAGCACCATGTCATGAACCTCTCCTACTACCCCACTTCCAGCAGAGGGGCA	399
591	-----GlyProGlyProAlaGlyAlaProGly	600
398	CATAGTGATGCTCGCGCCAAACACACTCCAGC-----	363
601	AspLysGlyGluAlaGlyProSerGlyProGlyProThrGlyAlaArgGlyAlaPro	620
362	-----CAAAGTTAGCAGGTGACCAAGAGCTGGCTTTCCGGT---GCCGCAGCA	312
621	GlyAspArgGlyGluAlaGlyProProGlyProAlaGlyPheAlaGlyProProGlyAla	640
311	GCGGCTCACCCACAGCTCT-----	291
641	AspGlyGlnProGlyAlaLysGlyGluProGlyAspThrGlyValLysGlyAspAlaGly	660
290	-----GGACCATAGTGGCCAGCGGG---TAGGGCTAGGGGGCGTTCAGGCATCC	240
661	ProProGlyPro-AlaGlyProAlaGlyProProGlyProLysGlyAsnValGlyAlaPr	680
239	AGAACTGCTCTGCTCTGCTGCCAGAGCTGCGGCCCTCTCCTCTTGTCTGCCGCCA	180
680	oGly-----ProLysGlyProArgGlyAlaAlaGlyProPro-GlyAlaThr-	695
179	ACTGCCTAGGAATCAGCCAGCGGCCCAATTTCTGCCAGCCCTTGGTCCGCGTCCAGCTTC	120
696	-----GlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSerGlyAsnA	713
119	TCAGCCCATGCTCAACACTGCTGCTGGGGCACCTCAGT-----GGGG	75
713	lacyProProGlyProProGlyProValGlyLysGluGlyProArgGlyG	733

1915 GTCAATACCTCTGTAGCAAGTAATGGCCAGCAGACCCAGGCTCGCGGCAGACAC 1856  
Db : : : : :  
107 ArgGlyPheSerGlyLeuAspGlyAlaLysGlyAlaGlyProAlaGlyProLysGly 126  
QY : : : : :  
1855 CATATAGCAGTGTACAGACTGGCTGAGCTGGACAAATGAGCCCAATAACAGGATGGGC 1796  
Db : : : : :  
127 GluProGlySerProGly- - - - -GluAsnGlyAlaProGlyGln--MetGlyP 142  
QY : : : : :  
1795 CA- - - - -CCTGGGACAGCAGGAGGCACTATCCAGGATGCGGAGGTCCAGGCAGAT 1745  
Db : : : : :  
142 roArgGlyLeuProGlyGluArgGlyArg- - - - -ProGlyAlaProGlyProAlaGlyA 160  
QY : : : : :  
1744 GCGCCGCGCCGAA- - - - -CCACCTGCGCTCGGTGGCTCACC 1706  
Db : : : : :  
160 laArgGlyAsnAspGlyAlaThrGlyAlaAlaGlyProGlyProThrGlyProAlaG 180  
QY : : : : :  
1705 CACCACCACA- - - - -CCTAGGAGACATCACAG- - - - - 1678  
Db : : : : :  
180 lyProGlyPheProGlyAlaValGlyAlaLysGlyGluAlaGlyProGlnGlyProA 200  
QY : : : : :  
1677 - - - - -GCAGAGCCCGCCAGAGCGGGTGGAGTGGAGCGCCCACTGCTCCAGCAC 1623  
Db : : : : :  
200 rgGlySerGluGlyProGlnGlyValArgGly- - - - -GluProGlyProProGlyP 217  
QY : : : : :  
1622 CCAGTGTCTTATAGG- - - - -AAGGAGCTCCAGGCTTAGG- - - - -CCTGGCAGGAAGC 1572  
Db : : : : :  
217 roAlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGlnProGlyAlaLys- 236  
QY : : : : :  
1571 TGGTCATCAGGCTGCTCACTGCTAGCACTCCAGTGTCCCTCGGTATTTGGGCAGGA 1512  
Db : : : : :  
237 GlyAlaAsnGlyAlaPro- - - - -Glylle- - - - -AlaGly 246  
QY : : : : :  
1511 ACACCTGCTTCTCCGGTGTAGAGG- - - - -AGGCCAGTGTGTAGGCCA 1467  
Db : : : : :  
247 AlaProGlyPheProGlyAlaArgGlyProSerGlyProGlnGlyProGlyGlyProThr 266  
QY : : : : :  
1466 GGATCTCGAGCGGTGAGAAGTGAACCCGGTGGAGCGGCTGAAGCTGTCCACCAGGCCA 1407  
Db : : : : :  
267 GlyAlaArgGlyLeu-ValGlyGluProGlyProAlaGlySerLysGlyGluSerGlyAs 286  
QY : : : : :  
1406 CACTGTGGGACAGCATGTGCCACCGCCAGCCACAGGA- - - - - 1368  
Db : : : : :  
286 nLys- - - - -GlyGluProGlySerAlaGly- - - - -ProGlnGlyProGlyProSerGlyGlu 304  
QY : : : : :  
1367 AAGCTGCCACACTGCCCAATAGACTGCTCGAGTGGCCGAATCGCTGCACCAGCGGTCCA 1308  
Db : : : : :  
305 GluGlyLysArgGlyProAsnGly- - - - -GluAlaGlySerAlaGlyProProGlyPro 322  
QY : : : : :  
1307 TGACCAGAGAACACAGGAGATGGCCACTGCAGGAACACGCCAGGCTGCCCATCC 1248  
Db : : : : :  
323 - - - - -ProGlyLeuArgGlyGly- - - - -ProGlySerArgGly 333  
QY : : : : :  
1247 GAACGCTTCATCATAGTGTCCGGGCTCGGTGCCGCTCAGCTCTGGGCACGCCCT 1188  
Db : : : : :  
334 PheProGlyAlaAspGlyValAlaGlyProLysGlyProAlaGlyGluArgGlySerPro 353  
QY : : : : :  
1187 GGTACAGCCCTCGCCCAACAGAAATCCGTGTAACACAGCGTGAAGGTCAATGAGTCCATCC 1128  
Db : : : : :  
354 GlyProAlaGlyProLysGlySerPro- - - - - 362  
QY : : : : :  
1127 ACCTGCACAGCTACGCCACAGAGCGCGGCAGGTTGGGGCATGCGGCACACAGCT 1068  
Db : : : : :  
363 : : : : :  
1067 GGTGACCGCGGGAAGCAGCGCCAGGT- - - - -TCCGGAAGCCACCGCG 1020  
Db : : : : :  
366 GlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGlySerProGlySer 385  
QY : : : : :  
1019 CCGGCGATGGACAGCAGTGGGGGACAAAGAGGGGGCCACAGCCCTTCTGCTGGCTCGG 960  
Db : : : : :  
386 ProGlyProAspGly- - - - -LysThrGlyProPro- - - - - 395  
QY : : : : :  
959 TGGGGCCCGCG- - - - -CTGCTCTCTCAGCCACCA- - - - - 930

396 - - - - -GlyProAlaGlyGlnAspGlyArgProGlyProProGlyProProGlyAlaArgGly 414  
QY : : : : :  
929 : : : : :  
415 GluAlaGlyValMetGlyPheProGlyProLysGlyAlaAla- - - - - 428  
QY : : : : :  
905 GGAAGATGAGGTTGAGCAGCAGGCCCAAGAGGACACT- - - - -CCTCCTGGTGGCCAGGT- - - 855  
Db : : : : :  
429 GlyGluProGlyLysAlaGlyGluArgGlyValProGlyProProGlyAlaValGlyPro 448  
QY : : : : :  
854 : : : : :  
449 AlaGlyLysAspGlyGluAlaGlyAlaGlnGlyProProGlyProAlaGlyProAlaGly 468  
QY : : : : :  
839 TGGTGTCCAGTCAATGGCAGCAGGA- - - - -GGTAGCCCA 804  
Db : : : : :  
469 GluArgGlyGluGlnGlyProAlaGlySerProGlyPheGlnGlyLeuProGlyProAla 488  
QY : : : : :  
803 GGCAGCCCCCAAGACTGATCATCAAGGCATAGACAGAGTAGCCCTGGCCAGACAGTGGT- - - 747  
Db : : : : :  
489 GlyProGlyGluAlaGlyLysProGlyGluGlnGlyValProGlyAspLeuGlyAla 508  
QY : : : : :  
746 CCGGGTCCC- - - - -GGAAGAGTGCAGACAGCAGGCGCTCCAGTGGAGTGAAGCACACT 593  
Db : : : : :  
509 ProGlyProSerGlyAlaArgGlyGluArg- - - - -GlyPheProGlyGlyAlaArgGlyValGln 527  
QY : : : : :  
692 GGCACAGAAGTCCAGCAGCCCGCCAGCGGATGAGCAGTGCCTCA- - - - - 648  
Db : : : : :  
528 GlyProGlyProAlaGlyProArgGlyAlaAsnGlyAlaProGlyAsnAspGlyAla 547  
QY : : : : :  
647 : : : : :  
548 LysGlyAspAlaGlyAlaProGlyAla- - - - -ProGlySerGlnGlyAlaProGlyLeu 565  
QY : : : : :  
602 CGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCAAGACACAGTGCACAGATGAAGG 543  
Db : : : : :  
566 GlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuProGly- - - - -ProLysGlyAsp 583  
QY : : : : :  
542 GCGGGGGGGCCATAGCTCCACGCCAGTGTGCTACTGGCTAGCCTAGGACG- - - - -GGA 486  
Db : : : : :  
584 ArgGlyAspAlaGlyProLysGlyAlaAspGly- - - - -SerProGlyLysAspGly 600  
QY : : : : :  
485 CACAGA- - - - -CCAGGCCCGCAGCTGGACCAATGCCCA- - - - - 453  
Db : : : : :  
601 ValArgGlyLeuThrGlyProIleGlyProGlyProAla- - - - -GlyAlaProGly 618  
QY : : : : :  
452 : : : : :  
619 AspLysGlyGluSerGlyProSerGlyProAlaGlyProThrGlyAlaArgGlyAlaPro 638  
QY : : : : :  
446 TGGTATGAACCTTCCTCTACCCCACTTCCAGCAGCAGAGGGCGGCACATAGTGTATGC 387  
Db : : : : :  
639 GlyAspArgGlyGluProGlyProGlyProAlaGlyPheAlaGly- - - - - 654  
QY : : : : :  
386 CTGCGGCCAAACACACACTCCAGCC- - - - -CAAGGTTAGCAGGTGACCAACA 339  
Db : : : : :  
655 : : : : :  
338 AGAGCTGGGCTTTCGGTTCGCCGAGCAGCGGCTCACCCACAGCTCTGGACCATAGTGG 279  
QY : : : : :  
670 AspAlaGlyAlaLysGlyAspAla- - - - -GlyProGlyProAlaGlyPro- - - - -AlaGly 687  
QY : : : : :  
278 GCCA- - - - -GGCGGTAGGCTCAGGGGGCGGTTTCAGGCATCTCCAGAACCTGCT 231  
Db : : : : :  
687 yProProGlyProIleGlyAsnValGlyAlaProGlyAlaLysGlyAla- - - - - 703  
QY : : : : :  
230 TCGTCTCGCTCTGCT- - - - -CCAGAAGCTCGCGGCTCTCTCTCTCTCTCTCTCTCT 177  
Db : : : : :  
704 - - - - -ArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgva 722  
QY : : : : :  
176 GCCTAGGAATCAGCCAGCGCGCCCATTTCTGCCAGCCCTTGGTGGCCGCTCAGCTTCTCA 117



```

185  CGCCCAACTGCCTAGAAATCAGCCAGGCGCCCATTC---TGCCAGCGCCTTGTGTGCC-- 131
Qy      |||      |||      |||      |||      |||      |||      |||
584  rArgGlyAspPro---GlyProProGlyAlaHisGlyProAlaGlyProLysGlyAlaHis 603
Db      |||      |||      |||      |||      |||      |||      |||
130  -GTGCCAGCTTCTCAGCCCATCTCAACACCTGCTGCTGTGGGCGACCTCAGTGGGGACA 72
Qy      |||      |||      |||      |||      |||      |||      |||
603  sGlyProAla-----GlyPr 608
Db      |||      |||      |||      |||      |||      |||      |||
71  CGTCTCATCTCAGATCCTGGCCGAGCGCGCGCTGTCTACCCGGAGCCAGCGCGTGA 12
Qy      |||      |||      |||      |||      |||      |||      |||
608  oLysGlyAlaMet-AspProGlyArg--TyrGlnLeuSerAlaGlyArgTyrHisTyrC 627
Db      |||      |||      |||      |||      |||      |||      |||
Qy      11  GCCTGGTT 4
Db      627  InLeuVal 629

Search completed: June 9, 2003, 22:35:21
Job time : 194.5 secs

```

Search completed: June 9, 2003, 22:35:21  
Job time : 194.5 secs

US-09-759-143-110 (1-3410) x Q9FVL6 (1-604)

Qy	193	GAGGAGAGCGCGCAGCTTCTGGAGCAGACCGAGACGAGTCTTGAGTGGCTCGAAC	252
Db	25	GluserArgPheAlaGlnLeuGluIleArgSerSerSerProArgValSerAsn	44
Qy	253	GGCCCCCTGAGCCCTACCCGCGCTGGCCCACTATGGTCCAGAGCGCTGGGTGAGCGCCT	312
Db	45	GlyCluMetAsnAspSerAsnLeu-----ProLeuProProProPro	58
Qy	313	GCTGGGCGACCGGAAAGCCAGCGCTCTGTGGTCAACCTGTGAACCTTTGGCGCTGAGGT	372
Db	59	-ValArgAsnSerLeuThrLeuIleLeuSerCysThrValAlaAalaglyValGlnPh	78
Qy	373	GTCTTTGGCGAGGCGATCACCTATCTGCGCGCTCTGCTGGAAGTGGGTGAGGCA	432
Db	78	eGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTyrIleGlnThrLeuGlyIleGluH	98
Qy	433	GAAGTTCATGACCATGGTCTGGCGATTTGGTCAGCTGGGCGCTGGTCTGTGTCGCCGT	492
Db	98	sAlaPheSerPheIleTrpLeuCysGlyProIleThrGlyLeuValValGlnProC	118
Qy	493	CTTAGGCTCAGGCAGTGACCACTGGCGTGGAGCGCTATGGCGCGCGCGCTTCATCTG	552
Db	118	sValGlyIleTrpSerAspLysCysHisSerLysTyrGlyArgArgProPheIlePh	138
Qy	553	GGCACTGTCTTGGGCATCCTGCTGAGCCCTTTCTCATCCCAAGGCGCGGTGGGTAGC	612
Db	138	eIleGlyAlaValMetIleSerIleAlaValIleIle-----GlyPheSerAl	155
Qy	613	A-----GGCGCTGTGCCGGAT-----CCCAG	636
Db	155	aAspIleGlyTrpLeuLeuGlyAspThrLysGluHisCysSerThrPheLysGlyThrAr	175
Qy	637	GCCCTGGAGCTGGCACTGCTCATCTCGGCGCTGGGCTGCTGGACTTCTGTGGCCAGGT	696
Db	175	gSerArgAlaAlaIleValPheValValGlyPheTrpMetLeuAspLeuAlaAsnAsnTh	195
Qy	697	GTGCTTCATCCACTGGAGGCCCTGCTCTGTGACCTCTCCGGGACCCGACCACTGTGC	756
Db	195	rValGlnGlyProAlaArgAlaLeuLeuAlaAspLeu---SerGlyProAspGlnArgAs	214
Qy	757	CCAGGCCCTACTCTCTATGCCCTCATGATCAGTCTTTGGGCGTGGCTGGGCTACCTCT	816
Db	214	nThrAlaAsnAlaValPheCysSerTrpMetAlaValGlyAsnIleLeuGlyPheSerAl	234
Qy	817	GCCTGCCATTGAC-----TGGGACACCACTGGCCCTGGCCCCCTACCTGGGCAC	864
Db	234	aglyAlaSerGlyGlyTrpHisArgTrp-----PheProPheLeuThrAs	249
Qy	865	CCAGGAGGATGTC-----CTCTTTGGCTGTGCTACCCCTCATCTTT	903
Db	249	nArgAlaCysCysGluProCysGlyAsnLeuLysAlaAlaPheLeuValAlaValPh	269
Qy	904	CCTCACC---TGGGTACAGCACACTGCTGGTGGTGAGGAGGACCGCTGGGCCCCAC	960
Db	269	eLeuThrLeuCysThrLeuValThrLeuTyrPheAlaAsnGluValProLeuSerProLy	289
Qy	961	CGAG---CCAGCAGAAGGGTGTGGCGCCCTCTCTTG---TCGCCCCAC	1003
Db	289	sGlnTyrLysArgMetSerAspSerAlaProLeuLeuAspSerProGlnAsnThrGlyPh	309
Qy	1004	---TGCTGTCCATCGCGGCGCGCTTGGCTTCCGAACCTG-----	1042
Db	309	eAspLeuSerGlnSerLysArgGluLeuGlnTyrArgAsnSerValAlaAsnAsnGluSe	329
Qy	1042	-----	1042
Db	329	rGluMetGlyHisValAlaAlaAspAsnSerProLysAsnGluGlnArgProAspLysAs	349
Qy	1043	-----GGCGCGCTGCTTCCCGCGCTGCACCAGCTGTG	1074

349	pGlnGlyAspSerPheAlaAspSerProGlyAlaValLeuValAsnLeuLeuThrSerLeu	369
1075	CTGCCGCATGCCCGCACCCCTGCGCGGCTCTTCGTGGCTGAGCTGTGCAGCTGATGGC	1134
369	uArgHisLeuProProAlaMethHisSerValLeuIleValMetAlaLeuThrTrpLeuSe	389
1135	ACTCATGACCTTCACCGCTGTTTACACGATTTTCGTGGCGGAGGCGCTCACCAGGCGCT	1194
389	rTrpPheProPhePheLeuPheAspThrAspTrpMetGlyArgGluValThrHisGlyAs	409
1195	GCCACGAGCTGAGCGGGACCGACCGCGGAGACACTATGATGAAGCGTTCCGGATGG	1254
409	pProLysGlyGlu-----AlaAspGluValAsnAlaTyAsnGlnGlyValArgGluG	427
1255	CAGCCTGGGCGCTGTTCTCGAGTCGCCCATCTCCCTGGTCTTCTCTGTCGTATGACCG	1314
427	yAlaPheGlyLeuLeuLeuAsnSerValValLeuGlyValSerPheLeuIleGluPr	447
1315	GCTGGTGCACGATTCGGCACTCGACAGCTCTATTGGCCAGT-----	1357
447	oMetCysLysTrpIleGlySerArgLeuValTrpAlaValSerAsnPheIleValPheVa	467
1358	-----GTGGCAGCTTTCCTGTGGCTGCTGCGGTCGCCGTCACATGCGCT	1395
467	lCysMetAlaCysThrAlaIleIleSerValValSerIleSerAlaAsnThrGlnGlyVa	487
1396	GTCCCAGCTGTGGCGGTGGTG-----ACAGCTTCAGCGCGCTCACCGGGTTCACCTT	1449
487	lGlnHisValIleGlyAlaThrArgSerThrGlnIleAlaAlaLeuValValPheSerLe	507
1450	CTCAGCCCTGCAGATC-----CTGCCCTACACACTGCCCTCCCTCTACCA	1494
507	uLeuGlyIleProLeuAlaValThrTyrsValProPheSerIleThrAla-----	524
1495	CCGGAGAACGAGTGTTCTGCCCCAATACCAGAGGGACACTGGAGGTCTAGCAGTGA	1554
525	-----GluLeuThrAlaAspAlaGly-----	531
1555	GGACAGCTGATGACCACTTCCTGCCAGCGCCCTAAGCCTGGAGCTCCCTTCCTAATGG	1614
531	-----	531
1615	ACAGCTGGTGCTGGAGGAGTGGCGTCTCCACCTCCACCGCGCTCTCGCGGGCGCTC	1674
532	-----GlyGlyGlnGlyLeu-----	536
1675	TGCCTGTGATGTCCTCCGTACGTGTGGTGGTGAGCCACCGAGCGAGGTGGTTC	1734
537	-----AlaIleGlyValLeuAsnLeuAlaIleValValPr	548
1735	GGGCGGGGATCTGCTCGACCTCGCCATCTCTGGATAGTGCCTTCCTGTGTGCCAGGT	1794
548	oGlnMetValValSerLeuGlyAlaGlyProTrpAspAlaLeuPheGlyGlyAsnII	568
1795	GGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTCTCACTGCCATAT	1854
568	e---ProAlaPheAlaLeuAlaSerLeuAlaAlaLeuAlaAlaGlyIlePheAlaMetLe	587
1855	G	1855
587	u	587

Search completed: June 9, 2003, 22:32:18  
Job time : 317.5 secs



[illegible]

Alignment Scores:	
Pred. No.:	5.05e-13
Score:	327.50
Percent Similarity:	37.29%
Best Local Similarity:	22.00%
Query Match:	5.10%
DB:	10
Length:	604
Matches:	141
Conservative:	98
Mismatches:	238
Indels:	164
Gaps:	21

Db 26 LeuPheLeuAlaCysMetValAlaGlyValGlnTyrGlyTrpAlaLeuGlnLeuSer 45  
 QY 395 TATGTGCCCTCTGCTGGAAGTGGGGTAGAGAGAGTTCATCACCATTGGTGTG 454  
 Db 46 LeuLeuThrProTyrValGlnThrLeuGlyLeuProHisAlaLeuThrSerValMetTrp 65  
 QY 455 GCAATTTGTCAGTCTGGGCTGTGTGTGTCCTCCAGGCTCAGCCAGTACAC 514  
 Db 66 LeuCysGlyProIleAlaGlyLeuIleValGlnProCysValGlyLeuTyrSerAspLys 85  
 QY 515 TGGCGTGGACGCTATGGCGCGCGCTTCATCTGGGCTGCTCTTGGGATCTG 574  
 Db 86 CysThrSerLeuGlyArgArgProPheIleLeuThrGlyCysIleIleCys 105  
 QY 575 CTGAGCTCTTCTCATC-----CCAGGGCGGCTGGCTAGCAGGCTGCTGTC 625  
 Db 106 IleSerValIleValIleGlyPheSerSerAspIleGlyTyrAlaLeuGlyAspThrThr 125  
 QY 626 CCGGAT-----CCAGGGCGGCTGGAGCTGGGCTGCTCATCTGCTGGC 667  
 Db 126 GluAspCysLysValTyrArgGlyProArgTyrHisAlaAlaAlaPheIleLeuGly 145  
 QY 668 GTGGGCTCTGCTGCTGCTGGCGAGGTGCTTCACTCCAGCTGGAGGCTGCTCTCT 727  
 Db 146 PheTrpLeuLeuAspPheSerAsnAsnThrValGlnGlyProAlaAlaLeuMetAla 165  
 QY 728 GACCTCTTCT-----CGGACCCCGGACCACTGTCGCCAGGCTACTGCTATGCCCTC 781  
 Db 166 AspLeuSerGlyArgHisGlyProSer-----AlaAlaAsnAlaIlePheCysSer 182  
 QY 782 ATGATCAGTCTGGGGCTGCTGGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTG 838  
 Db 183 TrpMetAlaLeuGlyAsnIleLeuGlyTyrSerSerGlySerThrAsnAspTrp----- 200  
 QY 839 AGTGCCCTGGCCCTTACCTGGCGGCGGCTGGCTTCTTCCGGAACCTGGC----- 877  
 Db 201 HisLysTrpPheProPheLeuMetThrArgAlaCysCysGluAlaCysAlaAsnLeuLys 220  
 QY 878 CTCTTTGGCTCTCACCTCATCTTCTCTC-----ACCTGGTAGCAGCCACACTGCTGTG 934  
 Db 221 AlaAlaPheLeuValAlaValPheLeuGlyLeuSerThrAlaValThrMetValPhe 240  
 QY 935 GCTGAGGAGGAGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 994  
 Db 241 AlaArgGluAlaLeuAspProValAlaAlaAlaLys----- 253  
 QY 995 TGCCGCCACTGCTGTCTCCTGCGGCGGCGGCTGGCTTCTTCCGGAACCTGGC----- 1045  
 Db 254 -----GCGCTGCTTCCCGGCTGTCACAGCTGTGCTGCGGATGCCCGGCGGCGGCTG 1096  
 QY 1097 CGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156  
 Db 281 -----IleValThrGlyLeuThrTrpLeuSerTrpPheProPheIleLeuPhe 296  
 QY 1157 TACACGGATTTCTGTCGGCGGCGGCTGTACCAGGCGGCTGCTGCTGCTGCTGCTG 1216  
 Db 297 AspThrAspTrpMetGlyArgGluIleTyrHisGlyArgProAspGlySerProAlaGlu 316  
 QY 1217 GAGGCGCGGAGACACTATGATGAGGCGGCTGGATGGGCTGGGCTGGGCTGCTGCTG 1276  
 Db 317 ValThrAla-----PheGlnGlyValArgGlnGlyAlaPheGlyLeuLeuLeuAsn 334  
 QY 1277 TGGCGCATCTCCCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1336  
 Db 335 SerIleValLeuGlyIleSerPheLeuIleGluProMetCysArgArgLeuGlyAla 354  
 QY 1337 CGAGCAGTCTATTGGCCAGCTGGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1396  
 Db 355 ArgAlaValTrpValMetSerSerAlaValValCysValAlaMetAlaAla----- 371

QY 1397 TCCACAGTGTGGCGTGTGACAGCTTACCGCGCTTACCGCGGTTACCTTCTCAGCC 1456  
 Db 372 -----ValSerValLeuSerAla----- 377  
 QY 1457 CTGAGATCTGCTGCTTACACACTGGGCTTCTTACCCGCGGAGAGCAGGTGTTCCTG 1516  
 Db 378 -----TrpSerLeu----- 380  
 QY 1517 CCCAAATACGAGGAGACACTGGAGGTGTAGCAGTACGAGACACCTGATGACACCTTC 1576  
 Db 381 -----GlyAspPheGlyGly-----SerValGlnAspAlaAlaArg----- 392  
 QY 1577 CTGCCAGGCGCTAAGCCTGAGCTGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1636  
 Db 393 -----AlaProAlaGluGlyGlyValArgAlaSerAlaLeu 405  
 QY 1637 GGCTGTGCTCCA-----CCTCCACCGCGCTGTGCGGGGCTCTGCTGTGTAT 1684  
 Db 406 AlaLeuPheValPheLeuGlyLeuProPheAlaValLeuCys----- 419  
 QY 1685 GTCTCCCTACGTGTGGTGGTGGAGCCACGAGGAGGCTGTTCGGGCGCGGCG 1744  
 Db 420 -----SerValProPheAlaValThrAlaGlnLeuAlaAlaSerArgGlyGlyGlnGly 438  
 QY 1745 ATCTGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1804  
 Db 439 LeuCys-----ThrGlyValLeuAsnIleSerIleValProGlnMetAlaAla 456  
 QY 1805 CTGTTTATGGCTCCATTTGCCAGTCT-----AGC 1834  
 Db 457 LeuGlyAlaGlyProTrpAspGluLeuPheGlyGluGlyAsnIleProAlaPheAlaMet 476  
 QY 1835 CAGTCTGCTACTGCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1882  
 Db 477 AlaSerValPheAlaAlaAlaAlaAlaAlaAlaGlyValValLeuLeu 492

## RESULT 44

Q9M535 PRELIMINARY; PRT; 530 AA.  
 ID AC Q9M535;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Sucrose transport protein.  
 OS Euphorbia esula (Leafy spurge).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Euphorbia.  
 OX NCBI\_TaxID=3993;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=UNDERGROUND ADVENTITIOUS BUDS;  
 RA Anderson J.V., Horvath D.P.;  
 RT "Identification of mRNAs expressed in underground adventitious buds of  
 RL Euphorbia esula (leafy spurge).";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF242307; AAF65765.1;  
 DR InterPro: IPR003662; sub transporter.  
 DR Pfam: PF00083; sugar\_tr\_1.  
 DR TIGRPFAMS: TIGR01301; GPH\_sucrose; 1.  
 SQ SEQUENCE 530 AA; 55843 MW; 059C1ED3BB02D356 CRC64;

## Alignment Scores:

Pred. No.: 4.94e-13 Length: 530  
 Score: 327.50 Matches: 137  
 Percent Similarity: 40.64% Conservative: 91  
 Best Local Similarity: 24.42% Mismatches: 212  
 Query Match: 5.10% Indels: 121  
 Db: 10 Gaps: 19

US-09-759-143-110 (1-3410) x Q9M535 (1-530)



2912 AGAAGAGGGGTGGTTAGGAAGCCCTTGAGACCTGAAGCCGCCACCCCTCTACCTTCCTT 2853  
162 GlyGluProGlyValSerGly---PrometGlyProArgGlyProGlyProGlyProGly 180  
2852 CAACACCCCTAACCTGGTAAACAGCATTTGGAAATTATCATTTGGGATAGTAGAATTCC 2793  
180 YLysPro----- 182  
2792 AAGTCTCTGGTTAGGCATTTTGGGGGCCAGACCCCGAGAGAGAGAGATTTCTGGCAATG 2733  
183 -----GlyAspSerGlyGluAlaGlyLys----- 190  
2732 ATCAGCCCAATGACACGATCTATCAGGGAGACCTGATTGTGGGGATCCGCCACCTACCC 2673  
191 -----ProGlyLysSerGlyGluArg-----GlyProGlyPro----- 202  
2672 AAATATTAGACCAACACAGAAAGCTAGCAATGATTCCTCTTACTTTGTTAAATAA 2613  
203 -----GlnGlyAlaArgGlyPheProGlyThrProGlyLeuPr 215  
2612 ATAGTTAAATATTAAATCCCTGTCTCTGTGTGATGGCAACAGAGAGACCAACAGGCCA 2553  
215 oGlyValLysGlyHisArgGlyThrProGlyLeuAspGlyAlaLysGlyGluAlaGlyAl 235  
2552 CATCCTGATAAAGGTAAAGAGGGGTGGATCAGCAAAAGACAGCTGCTGGGCTGAGG 2493  
235 aProGlyValLysGlySerGlySerProGlyGluAsnGlySerProGlyProMetG 255  
2492 GGACCTGGTTCTGTGTGGTGGCCCTCAGGACCTCTCCCTACAAATAAGTCATATGTC 2433  
255 y-----ProArgGlyLeu----- 259  
2432 AAATCCCATGAGGAGTGTTCATCTCTAGAACTCCCATGCAAGAGCTACATTAACGAA 2373  
260 -----ProGlyGluArgGlyArgThrGlyPr 268  
2372 GCTCAGGTAAAGGGCTTAGA-----GATGGAAACAGAGTACGATGATTTATTCAGC 2319  
268 oAlaGlyAlaGlyAlaGlyAlaArgGlyAsnAspGlyGlnProGly----- 282  
2318 TCCCAAAACCCCTCTCTAGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCCTGAGCCT 2259  
283 -ProAlaGlyProProGlyProValGlyProAlaGlyGly-----Pr 296  
2258 GGGTAATCCACCTCAGAGTCCCGCATCTCAGTGCATGGAGCCCTCTGCGCTCCCTGT 2199  
296 oGlyPheProGlyAlaProGlyAlaLysGlyGluAla---GlyProThrGly----- 312  
2198 ATAAGTCCAGACTGAACCCCTTGGAGGCTCCAGTCA----- 2159  
313 -----AlaArgGlyPro---GluGlyAlaGlnGlyProArgGlyGluProGlyTh 328  
2158 -----GGCAGCCCTAGAGCTGGGAGAGAGG----- 2132  
328 rProGlySerProGlyProAlaGlyAlaGlyAlaGlyAsnProGlyThrAspGlyIleProG 348  
2131 -----AGAGGAGCGGCC----- 2120  
348 yAlaLysGlySerAlaGlyAlaProGlyIleAlaGlyAlaProGlyPheProGlyProAr 368  
2119 -CAGCCCCCAGCTGGCAGCTACGCACCTCA-----GCACACAGGCTGGCAGCAGAGA 2067  
368 gGlyProProGlyProGlnGlyAlaThrGlyProLeuGlyProLysGlyGlnThrGlyG 388  
2066 GCCACATTACTTGGCAACAACAGAACTGGCGGCCAGCCCGGAGCCCTGAGGCTAA 2007  
388 uProGlyIleAlaGlyPheLysGlyGluGlnGlyProLysGlyGluProGlyProAlaG 408  
2006 CAGGAGCGGGAGCTGGGACCCAGT----- 1982  
408 yProGlnGlyAlaProGlyProAlaGlyGluGluGlyLysArgGlyAlaArgGlyGluPr 428  
1981 -----GAGGCGAGGCCCTCCACCCCAATGTCTGGAAGTTTCTACGCTGAGT 1935

428 oGlyGlyAlaGlyProValGlyProProGlyGluArgGlyAlaProGlyAsnArgGlyPh 448  
1934 ATTTGGCCCAAGTCTCTTCTC----- 1913  
448 eProGlyGlnAspGlyLeuAlaGlyProLysGlyAlaProGlyGluArgGlyProSerG 468  
1912 -----AAATACTACCTGTCTAGCAAAAGTAAATGGCGACCA----- 1878  
468 yLeuAlaGlyProLys-----GlyAlaAsnGlyAspProGlyArgProG 483  
1877 -GACCCAGCCCTGGCGGACAGACACCATATAGCAGCTGACAGCTGGCTGAGCTGACAAATG 1819  
483 yGluProGlyLeu-----ProGlyAlaArgGlyLeuThrGlyArgProGlyAspAl 500  
1818 GAGCCCAATAAAGGATGGGCGCA-----CCTGGGACAGCAGGAAG----- 1776  
500 aGlyProGlnGlyLysValGlyProSerGlyAlaProGlyGluAspGlyArgProGlyPr 520  
1775 -----CACTATCCAGATGGCGAGG 1756  
520 oProGlyProGlnGlyAlaArgGlyGlnProGlyValMetGlyPheProGlyProLysG 540  
1755 TCCAGCGCAGATGCCCGCG-----CGGAACACCCCTGGCCCTCG 1717  
540 yAlaAsnGlyGluProGlyLysAlaGlyGluLysGlyLeuPro-GlyAlaProGlyLeuA 560  
1716 GTGGGCTCACCCACCCACGACGACGACGACGAGAGGAGGCGCCCGCAGAGCGCG 1657  
560 rGlyLeu-ProGlyLysasp---GlyGluThrGlyAlaAlaGlyProGlyProAla 578  
1656 GTGGAGGTGGGAGGAGG----- 1639  
579 yProAlaGlyGluArgGlyGluGlnGlyAlaProGlyProSerGlyPheGlnGlyLeu 598  
1638 CCACGCTCCAGCACCCACGCTGCTCATTAGGGAAG----- 1603  
599 ProGlyProProGlyProGlyGlyGlyLysProGlyAspGlnGlyValProGly 618  
1602 -----GGAGCTCCAGGCTTA---GGGCTGGCAGGAAG-----TGCTCATC 1564  
619 GluAlaGlyAlaProGlyLeuValGlyPro-ArgGlyGluArgGlyPheProGlyGluAr 638  
1563 AGGCTGTCTCTACTGCTAGCAGCTCCAGTGTCCCTCGTATTGGCAGGAGAACACCTGC 1504  
638 gGlySerProGlyAlaGlnGlyLeuGlnGlyAlaArgGlyLeu---ProGlyThrProG 657  
1503 TTCTCCCGTGTGTAGAGGAGG-----CCAGTGTGTAGGCGAGGATCTGC 1459  
657 yThrAspGlyProLysGlyAlaSerGlyProAlaGlyProGlyAlaGlnGlyProPr 677  
1458 AGGCTGAGAAGGTGAACCCGCTGAGGCGGCTGAAGCTG-----TCACACCGCCAC 1406  
677 oGlyLeu-GlnGlyMetProGlyGluArgGlyAlaAlaGlyIleAlaGlyProLysGlyA 697  
1405 ACTGTGGACAGCATGTGCGACCGCACCGAGGAAAGCTGCCACACTGCGCAATAA 1346  
697 spArgGly---AspValGlyGluLysGlyProGlyAlaProGlyLysAspGlyGlyA 716  
1345 GACTGTCTGAGTGGCGGAATCGCT-----GCACACCGCTCCATGACACAGAGA---- 1296  
716 rGlyLeuThrGlyProLysGlyProGlyProAlaGlyAlaAsnGlyGluLysGlyG 736  
1295 -----AGACAGGAGGATGGCGCAGCTGCAGGACACCCCGCAGCTCCCGCCGAAAC 1244  
736 luValGlyProGlyProAlaGlyThrAlaGly---AlaArgGlyAlaProGlyGluA 755  
1243 GCCTTCATCATAGTGTCTCGGCGCTCGGTGCGGCTCAGCTCTGGCAGCGCTGTTA 1184  
755 rg-----GlyGluThrGlyProGlyProAlaGlyPheAlaGlyProGlyGlyA 772  
1183 CA----- 1182



```

QY 575 CTGAGGCTCTTCTCATC-----CAAAGGCGCGCTGGCTAGCAGGCTGCTGTC 625
Db ::::: :::::
QY 106 IleserValIleValIleValIleValIleValIleValIleValIleValIleVal 125
Db ::::: :::::
QY 626 CCGGAT-----CCAGGCGCGCTGGAGCTGGCACTCTCATCTCTGGC 667
Db ::::: :::::
QY 126 GluAspCysLysValTyrArgGlyProArgTyrHisAlaAlaAlaPheIleLeuGly 145
Db ::::: :::::
QY 668 GTGGGCTGCTGACTTCTGTGGCCAGGTGTCTCACTCCAGCTGGAGCGCTGCTCTCT 727
Db ::::: :::::
QY 146 PheTTPLeuAspPheSerAsnAsnThrValGlnGlyProAlaAlaArgAlaLeuMetAla 165
Db ::::: :::::
QY 728 GACCTCTTC-----CGGAGCCCGGACCACTGTGCCAGGCTACTCTGTCTATGCCCTTC 781
Db ::::: :::::
QY 166 AspLeuSerGlyArgHisGlyProSer-----AlaAlaAsnAlaIlePheCysSer 182
Db ::::: :::::
QY 782 ATGATCAGTCTTGGGGCTGCTGGCTACCTCTCTGCTGCCATC-----GACTGGGACACC 838
Db ::::: :::::
QY 183 TrpMetAlaLeuGlyAsnIleLeuGlyTyrSerSerGlySerThrAsnAspTrp----- 200
Db ::::: :::::
QY 839 AGTGGCCTGGCCCTTACTGTGGCCACCCAGGAGGAGTGC----- 877
Db ::::: :::::
QY 201 HisLysTTPheProPheLeuMetThrArgAlaCysGluAlaCysAlaAsnLeuLys 220
Db ::::: :::::
QY 878 CTCCTTGGCTGCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 934
Db ::::: :::::
QY 221 AlaAlaPheLeuValAlaValValPheLeuGlyLeuSerThrAlaValThrMetValPhe 240
Db ::::: :::::
QY 935 GCTGAGGAGGCGCTGGCCCGCCACCCAGGAGGCTGTGGCCCGCTCTCTCTCTCTCTCT 994
Db ::::: :::::
QY 241 AlaArgGluAlaLeuAspProValAlaAlaLys----- 253
Db ::::: :::::
QY 995 TCGCCCGCACTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1045
Db ::::: :::::
QY 1046 -----GCCCTGCTTCCCGGCTGCACCACTGTGCTCCCGGAGTCCCGCCCGCCAGCCCTG 1096
Db ::::: :::::
QY 261 GlyProLeuAlaValPheLysGlyMetLysAsnLeuProValGlyMetProSerValLeu 280
Db ::::: :::::
QY 1097 CGCCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
Db ::::: :::::
QY 281 -----IleValThrGlyLeuThrTrpLeuSerTrpPheProPheIleLeuPhe 296
Db ::::: :::::
QY 1157 TACAGGATTCCTGGGCGGCGGCTGTACAGGCGCTGTCCAGAGCTGAGCGCGGCACT 1216
Db ::::: :::::
QY 297 AspThrAspTrpMetGlyGluIleTyrHisGlyArgProAspGlySerProAlaGlu 316
Db ::::: :::::
QY 1217 GAGGCGCGGACACTATGATGAGGCGCTTGGATGGCGACCCCTGGGCTGCTCTCTCTCT 1276
Db ::::: :::::
QY 317 ValThrAla-----PheGlnGluGlyValArgGlnGlyAlaPheGlyLeuLeuAsn 334
Db ::::: :::::
QY 1277 TCGCGCATCTCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1336
Db ::::: :::::
QY 335 SerIleValLeuGlyIleSerSerPheLeuIleGluProMetCysArgArgLeuGlyAla 354
Db ::::: :::::
QY 1337 CGAGCAGTCTATTGGCCAGTGTGGCAGCTTTCCTGCTGTGGCTGCGGCTGCCACATGCTG 1396
Db ::::: :::::
QY 355 ArgAlaValTrpValMetSerSerAlaValValCysValAlaMetAlaAla----- 371
Db ::::: :::::
QY 1397 TCCACAGTGTGGCGTGGTGCAGCTTTCACCGCCCTCACCCTGCTCACTCTCTCACT 1456
Db ::::: :::::
QY 372 -----ValSerValLeuSerAla----- 377
Db ::::: :::::
QY 1457 CTGCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1516
Db ::::: :::::
QY 378 -----TrpSerLeu----- 380
Db ::::: :::::
QY 1517 CCCAAATCCGAGGAGCACATGGAGGTGCTAGCAGTGGAGACAGCCTGATGACCACTCTC 1576
Db ::::: :::::
QY 381 -----GlyAspPheGlyGly-----SerValGlnAspAlaAlaAtrg----- 392
Db ::::: :::::

```

```

QY 1577 CTGCCAGGCGCTAAGCCTGAGCTCCCTTCCCTAAATGGACACACGCTGGTGTCTGGAGCAGT 1636
Db ::::: :::::
QY 393 -----AlaProAlaGluGluGlyValArgAlaSerAlaLeu 405
Db ::::: :::::
QY 1537 GGCTGTCTCCCA-----CCTCCACCCGCGCTGTGCGGGGCTCTGCTGCTGTGAT 1684
Db ::::: :::::
QY 406 AlaLeuPheValPheLeuGlyLeuProPheAlaValLeuCys----- 419
Db ::::: :::::
QY 1685 GTCTCCCTACGTGTGTGGTGGTGAGCCACCCAGGAGGCTGGTTCCTGGGCGCGGCG 1744
Db ::::: :::::
QY 420 ---SerValProPheAlaValThrAlaGlnLeuThrAlaSerArgGlyGlyGlyGly 438
Db ::::: :::::
QY 1745 ATCTGCTGCGACCTGCGCATCTGATGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1804
Db ::::: :::::
QY 439 LeuCys-----ThrGlyValLeuAsnIleValProGlnMetAlaIleAla 456
Db ::::: :::::
QY 1805 CTGTTTATGGCTCCATTTGCCAGTCT-----AGC 1834
Db ::::: :::::
QY 457 LeuGlyAlaGlyProTrpAspGluLeuPheGlyGluGlyAsnIleProAlaPheAlaMet 476
Db ::::: :::::
QY 1835 CAGTCTGCTACTGCTATATATGTTCTGCGCGAGGCTGGTGTCTGCTC 1882
Db ::::: :::::
QY 477 AlaSerValPheAlaAlaAlaAlaAlaAlaGlyValValLeuLeu 492
Db ::::: :::::

RESULT 41
O80605 PRELIMINARY; PRT; 594 AA.
ID O80605;
AC O80605;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative sucrose/H+ symporter (Sucrose transporter).
GN T17M13.3 OR SUC3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA.
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RA "Arabidopsis thaliana chromosome II BAC T17M13 genomic sequence.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE FROM N.A.
RA Meyer S., Truernit E., Huemmer C., Besenbeck R., Stadler R., Sauer N.;
RA "AtSUC3: an unusual sucrose transporter from Arabidopsis.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AC004138; AAC32907.1; -.
DR EMBL; AJ289165; CAB92307.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 594 AA; 63972 MW; DF075A3601299DA0 CRC64;

Alignment Scores:
Pred. No.: 4,34e-13 Length: 594
Score: 328.50 Matches: 120
Percent Similarity: 38.63% Conservative: 72
Best Local Similarity: 24.14% Mismatches: 174
Query Match: 5.12% Indels: 131
DB: 10 Gaps: 18

US-09-759-143-110 (1-3410) x O80605 (1-594)
QY 347 AACCTGCTAACCTTTTGGCTGGAGGTGTGTTGGCGCAGCATCACCTAT----- 397
Db ::::: :::::
QY 59 SerLeuValThrLeuValLeuSerCysThrValAlaAlaGlyValGlnPheGlyTrpAla 78
Db ::::: :::::

```

Db 640 nGlyAlaThrGly--GluProGlyLysAlaGlyGluInGlyProGlyGluValGly 659  
 QY 1151 -----CGGTGAAGGTCATGAGTG 1134  
 Db 660 GlyProGlyProSerGlySerArgGlyAspArgGlyPheProGlyGluArgGlyGlyIle 679  
 QY 1133 CCATCCAGCTGCACAGCTCAGCCACGAGCCGCGGCGGAGTGCAGGCGATCGCGCAGC 1074  
 Db 680 GlySerAlaGlyProAlaGlyProArgGlyAlaAsnGlySerProGlyAsnAspGlyAla 699  
 QY 1073 -----ACAGCTGTGTCAGCGCGGGAAGCAGGCGCGCCAGGT---TCCGGA 1032  
 Db 700 ArgGlyGluSerGlyAlaAlaGlyAlaProGlyGlyMetGlyAlaProGlyLeuInGly 719  
 QY 1031 AAGCCAGC-----GGGCCCGCATCGACAGCAGTGGGCGGCAAGGAGGGCGCGACA 978  
 Db 720 MetProGlyGluArgGlySerGlyGlyAsnSerGlyAlaLysGlyGluArgGlyAspGly 739  
 QY 977 GCCCTTCTGCTGCTGGTGGGCGCCAGCGCTCCCTCCTCAGCCACAGCAGTGTGGCTG 918  
 Db 740 GlyProLysGlyAlaAspGlyGlyPro----- 748  
 QY 917 CTAGCGAGGTGAGAGATGAGGTGAGCAGGCGCAAGAGGCACTCTCTGGTCCCA 858  
 Db 749 -----GlyLysAspGlyMetArgGlyMetThrGly---ProLysGlyProPro 763  
 QY 857 GGTAGGGCGCAGGCGCACTGCTGCCAGTCAATGTCAGCAGGAGGTAGCCAGCAGCAGC 798  
 Db 764 GlyProThrGlyAlaHisGlyGluLysGlyGluGlyGlyLeuGlyProProGlyPro 783  
 QY 797 CCCCAGACTGATCATGAAGCATACAGAGTAGGCGCTGGCGACAGTGGTCCG----- 744  
 Db 784 ThrGlyGly-----ArgGlySerProGlyGluArgGlyGluHisGly 797  
 QY 743 GTTCCCGAGAGGTCAGACAGCAGCGCTCCAGTGGAGTGAAGCACACCTGCCACACA 684  
 Db 798 AlaProGlyProAlaGlyPheAlaGlyProGlyAlaAspGlyGlnProGlyAsnLys 817  
 QY 683 -----AGTCCAGCAGCGCCCGCCAGGATGAGCA 654  
 Db 818 GlyGluThrGlyAsnAsnGlyProLysGlyGluAlaGlyAlaProGlyProGlyGlyPro 837  
 QY 653 GTGCCAGCTCCAGGGGCTGGGATCCG---GGCACAGCAGCCCTGTAGCCAGCGCGCC 597  
 Db 838 ValGlyAlaProGlyProGlnGlyProAlaGlyAsnSerGlyThrLysGlyThrArgGly 857  
 QY 596 TTGGATGAGAAAGGCTCAGCAGGATGCCAAGGACAGTCCAGATGAAGGCGCGC 537  
 Db 858 AlaProGlyProProGlyAlaSerGlyMetPro-----Gly 869  
 QY 536 GCGGCGCATAGCTCCAGCGCAGTGTCTAGCTGAGCTAGGAGCGGACACAGCA 477  
 Db 870 ProGlyGlyArgValGlyProProGly-----GlySerGly---AlaPro 883  
 QY 476 GCGCCAGCATGGACCAATGCCAGCAGCATGTGTCATGAACCTCTCTACCCCACTT 417  
 Db 884 Gly-----SerAlaGlyProProGly 890  
 QY 416 CCAGCAGCAGAGCGGCACATAGTGTATGCTGCGGCCAAACACACACTCCAGCGCAAGG 357  
 Db 891 ProAlaGlyLysGluGlyGlnLys-----GlyGlyArg 901  
 QY 356 -----TTAGAGGTGTACCAAGAGAGCTGGGCTTCCCGT----- 321  
 Db 902 GlyGluThrGlyIleAlaGlyArgProGlyGluAlaGlyAlaAlaProGlyProProGlyPro 921  
 QY 320 ---GCCCGCAGCGC-----GGCTCACCCACAGCCTCTGGA 288  
 Db 922 SerGlyAlaSerGlyAlaLysGlyAsnAspGlyProMetGlyAlaProGlyThrProGly 941  
 QY 287 CCATAGT-----GGCCAGGCGGGTAGGCGCTCAGG 258  
 Db 942 Pro-GlyGlyIleAlaGlyGlnArgGlyIleValGlyGlyProGlyGlyArgGlyProse 961

QY 257 GGCGCGTTCAGCACTCCAGAACTGCTCGTCTCGCTCTGCTCCAGAAAGCTGCGGCCTC 198  
 Db 961 rGlyThrAlaGlyIlePro-----GlyThrAlaGlyGluProGlyLysGly 976  
 QY 197 TCCTCCTTGTCTGCCGCCCAACTGCTAGGAATCAGCCAGCGGCCCTTTCTGCCAGCCCTT 138  
 Db 976 nGlyProGlyGlyProValGlyGluArgGlySerProGly-----ProMe 991  
 QY 137 TGGTCCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGGGGCAGCTCAGTG 78  
 Db 991 tGly-ProProGlyLeu-----SerGlyAlaProGlyGluAlaGlyArg---GluG 1007  
 QY 77 GGCACAGCTCTCATCTCAGTCTGCGCGA----- 46  
 Db 1007 lySerThrGlyHisAspGlyValSerGlyArgAspGlyProGlyProLysGlyAspA 1027  
 QY 45 -----GGCGCGGCGTGTCCCGGAGCC 22  
 Db 1027 rGlyGluAsnGlyAsnAlaGlySerProGlyAla 1038

## RESULT 40

Q944W2

ID Q944W2 PRELIMINARY; PRT; 506 AA.

AC Q944W2;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Sucrose transporter.

OS Oryza sativa (indica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI\_TaxID=39946;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. IR 36;

RA Whitfield P.R., Aoki N., Scofield G.N., Hirose T., Terao T.,

RA Furbank R.T.;

RT "Isolation and characterisation of a putative sucrose transporter gene

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF419298; AAL14982.1; -.

DR InterPro; IPR003662; sub.transporter.

DR Pfam; PF00083; sugar\_tr; 1.

DR TIGRFAMs; TIGR01301; GPH\_sucrose; 1.

SQ SEQUENCE 506 AA; 52774 MW; A4E73029D7022B64 CRC64;

## Alignment Scores:

Pred. No.: 4,23e-13 Length: 506  
 Score: 328.50 Matches: 143  
 Percent Similarity: 39.57% Conservative: 77  
 Best Local Similarity: 25.72% Mismatches: 207  
 Query Match: 5.12% Indels: 129  
 DB: 10 Gaps: 21

US-09-759-143-110 (1-3410) x Q944W2 (1-506)

QY 335 CTCTTGTGCTCAACCTGCTTAACCTTTGGCTGGAGGTGTGTTGGCCGAGCATCAC 394  
 Db 26 LeuPheLeuAlaCysMetValAlaGlyGlyValGlnTyrGlyTrpAlaLeuGlnLeuSer 45  
 QY 395 TATGTCCCGCTCTGCTGCTGGAGTGGGGGTAGAGAGAAGTTCTATGACCATGGTGCTG 454  
 Db 46 LeuLeuThrProTyrIleGlnThrLeuGlyIleProHisAlaLeuThrSerValMetTrp 65  
 QY 455 GGCATGTGTCCAGTCTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 514  
 Db 66 LeuCysGlyProIleAlaGlyLeuIleValGlnProCysValGlyLeuTyrSerAspLys 85  
 QY 515 TGGCGTGGACGCTATGGCGCGCGCGCTTCATCTGCGGCACATGCTGCTGCTGCTGCTGCTG 574  
 Db 86 CysThrSerSerLeuGlyArgArgProPheIleLeuThrGlyCysIleIleLeuCys 105







Db 531 yGluArgGlyProSerGlyLeuAlaGlyProLys-----GlyAlaAsnGI 546  
 QY 1883 CGACCA-----GACCAGGCTCGGGCAGACACATAGCAGTGCAGAC 1837  
 Db 546 yAspProGlyArgProGlyGluProGlyLeu-----ProGlyAlaArgGlyLeuTh 563  
 QY 1836 TGGCTGAGCTGACAAATGAGGCCATAACAGGGATGGGCCA-----CCTGGGAC 1786  
 Db 563 rGlyArgProGlyAspAlaGlyProGlnGlyLysValGlyProSerGlyAlaProGlyGI 583  
 QY 1785 AGCAGGAGG-----CA 1774  
 Db 583 uAspGlyArgProGlyProGlyProGlyProGlnGlyAlaArgGlyGlnProGlyValMetGI 603  
 QY 1773 CTATCCAGATGGCAGGTCCAGGAGATGCCCGCC-----CC 1735  
 Db 603 yPheProGlyProLysGlyAlaAsnGlyGluProGlyLysAlaGlyGluLysGlyLeuPr 623  
 QY 1734 GGAACCCCTGGCTCGTGGGTCCACCCACCCACCATAGGAGACATCAGGCA 1675  
 Db 623 o-GlyAlaProGlyLeuArgGlyLeu-ProGlyLysAsp---GlyGluThrGlyAlaAla 641  
 QY 1674 GAGGCCCGCAGAGCGGTGGAGGTGGAGCAGG----- 1639  
 Db 642 GlyProGlyProAlaGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyAlaProGlyPro 661  
 QY 1638 -----CCATGCCCTCCAGACCCACGTGTCCATTAGGGAAG----- 1603  
 Db 662 SerGlyPheGlnGlyLeuProGlyProGlyProGlyGluGlyGlyLysProGly 681  
 QY 1602 -----CGAGCTCCAGGCTTA-----GGGCTCGGCAAGC-- 1572  
 Db 682 AspGlnGlyValProGlyGluAlaGlyAlaProGlyLeuValGlyPro-ArgGlyGluAr 701  
 QY 1571 -----TGGTCATCAGGCTGTCTCTACGTGTAGCACCTCCAGTGTCCCTCGGTAT 1522  
 Db 701 gGlyPheProGlyGluArgGlySerProGlyAlaGlnGlyLeuGlnGlyProArgGlyLe 721  
 QY 1521 TTGGCCAGACACTGCTTCTCCCGTGTAGGGAGG-----CCAGT 1477  
 Db 721 u-----ProGlyThrProGlyThrAspGlyProLysGlyAlaSerGlyProAlaGlyProPr 740  
 QY 1476 GTGTAGGCGAGATCTGCAGGCTGAGAAAGTGAACCGGTGAGGGCGGTGAGCTG-- 1419  
 Db 740 oGlyAlaGlnGlyProProGlyLeu-GlnGlyMetProGlyGluArgGlyAlaAlaGlyI 760  
 QY 1418 -----TCACACGGCCACACTGTGGACAGCATGTGGCCACCGCCAGCCACAGGAAAGC 1364  
 Db 760 leAlaGlyProLysGlyAspArgGly---AspValGlyGluLysGlyProGluGlyAlaP 779  
 QY 1363 TGCCACACTGCCCAATAGACTGCTCGNGTCCCAATCGCT-----GCACCGCCGGTC 1310  
 Db 779 roGlyLysAspGlyGlyArgGlyLeuThrGlyProLysGlyProGlyProAlaGlyA 799  
 QY 1309 CATGACCAGAGAGA-----AGACGAGGAGATGGCGCACTGCAGGACAGCC 1262  
 Db 799 laAsnGlyGluLysGlyGluValGlyProProGlyProAlaGlyThrAlaGly---AlaA 818  
 QY 1261 CAGGCTGCCATCCAGACGCTTCATAGTGTCTCCGGGCTCGGTGCCCGCTCAGC 1202  
 Db 818 rGlyAlaProGlyGluArg-----GlyGluThrGlyProProGlyProAlaGlyP 835  
 QY 1201 TCTGGGCGCCCTGGTACA----- 1182  
 Db 835 heAlaGlyProProGlyAlaAspGlyGlnProGlyAlaLysGlyGluGlnGlyGluAlaG 855  
 QY 1181 -----GCCCTCGCCACGAAATCCGTGTATAAACAGCGTGAAGT 1142  
 Db 855 LyGlnLysGlyAspAlaGlyAlaProGlyProGlnGlyPro---SerGlyAlaProGlyP 874  
 QY 1141 CATGAGTCCATCCAGCTGCAGCTCAGCCACAGAGCGCGCAGGTGCGGGCAT 1082  
 Db 874 roGlnGlyProThrGlyValThrGly---ProLysGlyAlaArgGlyAlaGlnGlyProp 893

QY 1081 GCGCAGCACAGCT-----GGTGCAGCCCGGGAGCAGAGGCGCCCGCAGGTTC--GGAA 1031  
 Db 893 roGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProProGlySerAsnGlyA 913  
 QY 1030 AGCCAAGCGCGCGCGCATG----- 1011  
 Db 913 snProGlyProProGlyProProGlyProSerGlyLysAspGlyProLysGlyAlaArgG 933  
 QY 1010 --CACACAGTGGGGCACAAGGAGGGCGGACAGCCCTTCGTGCTGCTGGGTGGGGC 953  
 Db 933 lyAspSerGlyProProGlyArgAlaGlyAspProGlyLeu-----GlnGlyP 949  
 QY 952 CAGCGTGCCTCT----- 938  
 Db 949 roAlaGlyProProGlyGluLysGlyGluProGlyAspGlyProSerGlyProAspG 969  
 QY 937 AGCCACAGCAGTGTGGCTGTACGAGGTGAGGAAGATGAGGCTGAGCAGGAGGAGGAGG 878  
 Db 969 lyProPro-----GlyProGlnGlyLeuAlaGlyGlnArg 981  
 QY 877 GCACCTCT-----CCTGGGTGCCAGGT----- 855  
 Db 981 lyIleValGlyLeuProGlyGlnArgGlyGluArgGlyPheProGlyLeuProGlyProS 1001  
 QY 854 -----AGGGG----- 849  
 Db 1001 erGlyGluProGlyLysGlnGlyAlaProGlyAlaSerGlyAspArgGlyProProGlyP 1021  
 QY 848 -----CCAGGCGACTGTGTGCCAGTCAATGGCAGGAGGAGGAGGAGGAGGAGG 800  
 Db 1021 roValGlyProProGlyLeuThrGlyProSer-----GlyGluProGly- 1035  
 QY 799 GCCCCCCAAGACTCATGAGGCATAGACAGTAGTAGGCTGGCGACAGTGGTCCGGTGC 740  
 Db 1036 -----ArgGluGlySerProGlyAlaAspGlyPro---P 1046  
 QY 739 CCGAAGAGGTCTCAGAGCAGGCG-----CCTCCAGTGGAGT 704  
 Db 1046 roGlyArgAspGlyAlaAlaGlyValLysGlyAspArgGlyGluThrGlyProAl- 1064  
 QY 703 GAAGCAGCTGCCACAGAAAGTCCAGCAGCCCA-----CGCCAGGATGAGCAGTCCAG 647  
 Db 1065 --GlyAlaProGly-----AlaProGlySerProGlySerProGlyProAlaGlyProT 1082  
 QY 646 CTCAGGGGCTGGGATCCGGCAGCAGCAGCTGTAGCCAGCGCGCTTGGGATGAG 587  
 Db 1082 hrGlyLysGlnGlyAspArgGlyGluAlaGly-----AlaGlnGlyProMetGly---P 1099  
 QY 586 AAGAGGCTCAGCAGGATGCCAAGGACAGTGCCT-----AGATGAAGGGCGCGG 536  
 Db 1099 roAlaGlyProAlaGlyAla---ArgGlyIleProGlyProGlnGlyProArgGlyAspL 1118  
 QY 535 GCGGCCATAGCTGCCAGCAGTGGTCTGCTGAGCTGAGCTAGGCGGAGGAGCAGACA- 480  
 Db 1118 ysGlyGluAlaGlyGluAlaGlyGluArgGlyLys-----GlyHisArgGlyP 1135  
 QY 479 -----CCAGGCGCAGCAGTGGACCAA-----T 458  
 Db 1135 hethrGlyLeuGlnGlyLeuProGlyProProGlyProSerGlyAspGlnGlyAlaSerG 1155  
 QY 457 GCGCAGCA---CCATGGTCACTGAACTTCTCTACCCCACTT---CCAGCAGCAGAGG 404  
 Db 1155 lyProAlaGlyProSerGlyProArgGlyProProGlyProValGlyProSerGlyLysA 1175  
 QY 403 CGGCACATAGGTGATCCCTGCGGCCAACAACACACCTCCAGGCCCAAGGTTAGCAGGTTGAC 344  
 Db 1175 spGlyAlaAsnGlyIleProGlyProIleGlyProProGlyProGly----- 1190  
 QY 343 CAGCAAGAGCTGGGCTTTCGGT---GCCGACAGGCGGCTCAGCCACAGCCTCTGAC 287  
 Db 1191 --GlyArgSerGlyGluThrGlyProAlaGlyProGlyAsnProGlyProProGlyP 1210

```

Db 454 SerValGlyGlyGlyProPheAspGluLeuPheHisGlyGlyAsnIleProAla----- 471
Qy 1592 CTTGGAGCTCCCTCCCTATGACACAGCTGGGTGCTGGAGGAGTGGCTGCTC----- 1645
Db 472 -----PheValLeuGlyAlaIleAlaAlaValSerGlyIleLeuAlaLeu 487
Qy 1646 -----CCACCTCCACCGCG 1660
Db 488 ThrValLeuProSerProProAspAla 497

RESULT 37
ID 077753 PRELIMINARY; PRT; 1487 AA.
AC 077753;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Type IIA procollagen.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98340920; PubMed=9676231;
RA Du F., Acland G.M., Ray J.;
RT "Differential splicing of type II procollagen mRNA in canine retina.";
RL AnIm. Biotechnol. 9:15-20(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20480698; PubMed=11024291;
RA Du F., Acland G.M., Ray J.;
RT "Cloning and expression of type II collagen mRNA: evaluation as a
RT candidate for canine oculo-skeletal dysplasia.";
RL Gene 255:307-316(2000).
DR EMBL; AF023169; AAC62178.2; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 18.
DR Pfam; PF00093; vwc; 1.
DR ProDom; PD000007; Collagen; 4.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWF; 1.
KW Collagen.
SQ SEQUENCE 1487 AA; 141875 MW; 25873EAC1E311DB8 CRC64;

Alignment Scores:
Pred. No.: 2,79e-13 Length: 1487
Score: 332.50 Matches: 315
Percent Similarity: 31.42% Conservative: 62
Best Local Similarity: 26.25% Mismatches: 414
Query Match: 5,34% Indels: 411
DB: 6 Gaps: 67

US-09-759-143-110 (1-3410) x 077753 (1-1487)
Qy 3032 GTGGGGAAGTTGGGGTAGGGAAAGTTGGGGTAGGGGAAATTTTGGGAGTGCCTTC 2973
Db 203 MetGlyProMetGlyProArgGlyProProGlyProAlaGlyAlaProGlyPro----- 220
Qy 2972 ATCAGCCAGTCCCTAGAGAGATAGACGGGAGTGGAGTGGGGGGAACAGGCTGGGCCA 2913
Db 221 -----GlnGlyPheGlnGlyAsnProGlyGluPro 230
Qy 2912 AGAGAAGAGGGTGTAGGAAGCGCTTGAGACTGAAGCCCAACCTCTACCTCTCTT 2853
Db 231 GlyGluProGlyValSerGly---ProMetGlyProArgGlyProProGly-ProProGly 249
Qy 2852 CAACACCCTAACCTTTGGGTAAACAGCAATTTGGAATTCATTTGGGATGAGTAGAATTCC 2793

```

```

Db 249 yLysPro----- 251
Qy 2792 AAGGTCCTGGGTTAGGCATTTTGGGGGCCAGACCCAGGAGAGAAGATTCTGCAATG 2733
Db 252 -----GlyAspGlyGluAlaGlyLys----- 259
Qy 2732 ATCAGCCCAATGACCACTATCTCAGGGGACCTGATTGTGGGATCCGCCACCTACCC 2673
Db 260 -----ProGlyLysSerGlyGluArg-----GlyProProGlyPro----- 271
Qy 2672 AAATATTAGACACACACAGAAAAGCTAGCAATGGATTCCCTTCTACTTTCTTAAATAA 2613
Db 272 -----GlnGlyAlaArgGlyPheProGlyThrProGlyLeuPr 284
Qy 2612 ATAAGTTAAATATTAAATGCCTCTCTCTCTGATGGCAACAGAGGACCAACAGGCCA 2553
Db 284 oGlyValLysGlyHisArgGlyTyProGlyLeuAspGlyAlaLysGlyGluAlaGlyAl 304
Qy 2552 CATCTGTATAAAAGGTAAAGGTGATCAGCAAAAAGACAGTGTCTGTGGCTGAGG 2493
Db 304 aProGlyValLysGlyGluSerGlySerProGlyGluAsnGlySerProGlyProMetG 324
Qy 2492 GGACCTGGTCTGTGTGTGTGTGGCCCTCAGGACTCTTCCCTACAAATAAGTCATATGTC 2433
Db 324 y-----ProArgGlyLeu----- 328
Qy 2432 AAATCCCATGGAGAGTGTTCATCCTCAGAACTCCCATGCAAGAGCTACATTAAACGAA 2373
Db 329 -----ProGlyGluArgGlyArgThrGlyPr 337
Qy 2372 GCTGCAGTTAAGGGCTTAGA-----GATGGAAACAGGTGACTGAGTTATTATTCAGC 2319
Db 337 oAlaGlyAlaAlaGlyAlaArgGlyAsnAspGlyGlnProGly----- 351
Qy 2318 TCCCAAAACCTTCTCTAGTGTGTCTCACTAGGAGGCTAGCTGTTA----- 2270
Db 352 -ProAlaGlyProProGlyProValGlyProAlaGlyGlyProGlyPheProGlyAlaPr 371
Qy 2269 -----ACCCTGAGCCTGGGTAACTCCACCTCCAGAGTCCCG 2235
Db 371 oGlyAlaLysGlyGluAlaGlyProThrGlyAlaArgGlyProGluGlyAlaGlnGlyPr 391
Qy 2234 GCATTCAGTGCATGAGCCCT---TCTGGCCTCCTCTGATATAGTCCAGCTGAAACCCC 2178
Db 391 oArgGlyGluProGlyThrProGlySerProGlyProAlaGlyAlaSerGlyAsnProG 411
Qy 2177 CTTGGAGGCCTC-----CAGTCAGCAGCCCTTAGAGACTGGGAGAGAG 2133
Db 411 yThrAspGlyIleProGlyAlaLysGlySerAlaGlyAlaProGlyIleAlaGlyAlaPr 431
Qy 2132 G-----AGAGGAGCCGCCAGCCCTGTGTGACGTACGACCTCA-----GCAGC 2085
Db 431 oGlyPheProGlyProArgGlyProProGlyProGlnGlyAlaThrGlyProLeuGlyPr 451
Qy 2084 ACAGGTGGCAGCAGAGAGCCACATTACTTTTGGCAGCAACAGAAACTGGGGCCAGCCCG 2025
Db 451 oLysGlyGlnThrGlyGluProGlyIleAlaGlyPheLysGlyGluGlnGlyProLysG 471
Qy 2024 GCAGCCCATGGGCTTAACAGGAGCGGGAGTGGGACCCAGT----- 1982
Db 471 yGluProGlyProAlaGlyProGlnGlyAlaProGlyProAlaGlyGluGlnGlyLysAr 491
Qy 1981 -----GAGGAGGAGCCCTCCACCCCAATGTGCTGG 1953
Db 491 gGlyAlaArgGlyGluProGlyAlaGlyProValGlyProGlyProGlyGluArgGlyAl 511
Qy 1952 AAGTTTCTAGCTGAGTATTGGCAAGTCGCTCTTGTG----- 1913
Db 511 aProGlyAsnArgGlyPheProGlyGlnAspGlyLeuAlaGlyProLysGlyAlaProG 531
Qy 1912 -----AAATACTACCTGTGTAGCAAGTAATGG 1884

```



QY	2175	TGGAAGGCTTCAGTCAGCGAGCCCTAGAGACTGGGGAGAG-----AGGA	2131
Db	971	ysaspelyLeu---ProGlyHisProGlyGlnArgGlyGluThrGlyPheGlnGlyLysT	990
QY	2130	GAGGAGCGCCCGCCAGCCCTGTCAGCTACCCACTCAGCAGCAGGCTGGCAGCA	2071
Db	990	hrGlyProProGlyProGlyVal-----ValGlyProGlnGlyProThrG	1006
QY	2070	GAGGCGCACATTACTTGGCAGCAACAGAAACTGGCGCCAGCCCGCAGCCCC-----	2017
Db	1006	lyGluThrClyPrometGlyGluArgGlyHis-ProGlyProProGlyProGlyGlu	1025
QY	2016	ATGGGGCTTAACAGCGGGGAGCTGGACCCACTGAGGAGCGCCCTCCACCCCAATGTG	1957
Db	1026	GlnGlyLeuProGlyLeu-Ala-----GlyLysGluGlyThr-----	1037
QY	1956	CTGGAGATTTCCTACGCTGAGTATTGGCCAAAGTCGCTCTGTGCAATACCTCTGTGA	1897
Db	1037	-----	1037
QY	1896	GCAAGTAATAATGGCAGCAGACACCCAGCGCTCGGCAGACACCAATATAGGCAGTGACAGAC	1837
Db	1038	-----LysGlyAspProGlyProAlaGlyLeuProGlyLysAsp-----	1050
QY	1836	TGGCTGAGCTGACAACTGGAGCCCATAAACAGGGATGGGCCACTCGGACAGCAGGAG	1777
Db	1051	-----GlyProProGlyLeuArgGly--	1057
QY	1776	GCACATATCAGGATGGCAGGTCCAGGCAGATGCCCGGC-----	1737
Db	1058	----PheProGlyAspArgGly-----LeuProGlyProValGlyAlaLeuGlyLe	1073
QY	1736	-----CCGGACACACCCTGCGCTCGGTGGCTCACCACCACCCACACA	1696
Db	1073	uLysGlyAsnGluGlyProProGlyProProGlyPro--AlaGlySerPro-----	1089
QY	1695	CGTACGAGACATATCAGCAGAGCGCCCGCAGAGCGCGGTGGAGTGGGACAGCGCCA	1636
Db	1090	---GlyGluArgGlyProAlaGlyAlaAlaGlyProIleGlyIleProGly---ArgPro	1107
QY	1635	CTGCCTCCAGCACCCAGCTGCTCATTAGGG-----AAGGGAGCTCCAGGC-----	1591
Db	1108	GlyProGlnGlyProProGlyProAlaGlyGluLysGlyAlaProGlyGluLysGlyPro	1127
QY	1590	TTAGGCTTCGGCAGGAAGCTGCTCATCAGGTGTCTCTCACTGCTAGCACCTCTCAGTGTCC	1531
Db	1128	GlnGlyPro-AlaGly-----ArgAspGly-----LeuGlnGlyPr	1139
QY	1530	CTTCGGTATTGGGCAGGAACACCTGCTTCTCCGGTGGTAGAGGAGCGCCAGTGTGTAG	1471
Db	1139	oValGlyLeu-----ProGlyProAlaGlyProValGly-----ProProGlyG1	1154
QY	1470	GCAGGATCTGAGGCTCAGAAGGTGACCCGGTGGCGGGCTGAAGCTGTCCACCACG	1411
Db	1154	uaspglyAsp-LysGlyGluIleGlyGluProGlyGlnGlySerLysGlyAspLysG	1174
QY	1410	GCCAC-----ACTGTGGAGCAGCATGTGGCA	1384
Db	1174	lyGluGlnGlyProProGlyProThrGlyProGlnGlyProIleGlyGlnProGlyProS	1194
QY	1383	CGGCGACCCACAGGGAAGCTGCCACACTGGCCAAATAGACTGTCTCAGTCCCAATCGC	1324
Db	1194	erGlyAlaAspGlyGluProGlyProArgGlyGln-----	1205
QY	1323	TGCACACCGCGTCTCCATCCAGAGAGAAGACACGAGAGATGGC-----	1280
Db	1206	-----GlnGlyLeuPheGlyGlnLysGlyAspGluGlyProArgGlyPheProGlyProp	1224
QY	1279	-----GCATGTCAGGAACAGCCCGAGGCTGCCATCCGGAACCCCTTCATCATAG	1231
Db	1224	roGlyProValGlyLeuGln--GlyLeuProGlyProProGlyGluLysGlyGluThrG1	1243
QY	1230	TGTCTCCGGGCTCGGTGCCGCTCAGCTCTGGGACACGCCCTGTGTACAGCCCTCTCGGCC	1171
Db	1243	yasValGlyGlnMetGlyPro-----ProGly-----	1256
QY	1170	ACGAATCCCTGTAAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCC	1111
Db	1256	oArgGlyPro-----SerGlyAlaProGlyAlaAspGlyProGlnG1	1270
QY	1110	ACGAAGAGCCGGCGCAGGG---TGCGGGCGATGCGGCACGACAGCT-----GGTGCAGC	1060
Db	1270	yProProGlyGlyIleGlyAsnProGlyAlaValGlyGluLysGlyGluProGlyGluAl	1290
QY	1059	CGGCGAAGCAGGCGCCAGTTCCGGAAGCAACCGCGCGCCGCGCATGGACAGCAGTGG	1000
Db	1290	aglyGluProGlyLeuProGlyGluGlyProProGlyProLysGlyGluArgGlyG1	1310
QY	999	GGCGACAAAGAGGGGGCGCAGACCCCTTCTGCTGGCTCGGTGGGCGCCAGCCTCCCTCC	940
Db	1310	uLysGlyLeuSerGlyProSerGlyAlaAla-----GlyProProGlyProLy	1322
QY	939	TCAGCCACAGCAGCTGTGGCTGTACCGAGTCTAGGAAGATGAGGTGAGCAGCGCAAG	880
Db	1326	sglyProPro-----GlyAspAspGlyProLy	1335
QY	879	AGGCATCTCT-----CCTGGTCCCGAGGTAGGGCGCCAGGCACCTGTGTG	835
Db	1335	sglySerProGlyProValGlyPheProGlyAspProGly--ProProGlyGluProG1	1354
QY	834	TCCAGATCAATGGCAGCAGGAGTAGCCAGCAGCCGCCCAAGACTGATCATGAAGCA	775
Db	1354	yProAla-----GlyGlnAspGlyProProGlyAsp-----LysG1	1366
QY	774	TAGACAGAGTAGGCTGGCGACAGTGT---CCGGTCCCGGAAGAGGTCTCAGAGACAGG	718
Db	1366	yasAspGlyGluProGlyGlnThrGlySerProGlyPro-----ThrG1	1381
QY	717	GCCTCCAGTGGAGTGAAGCACACCTGCCACAGAGTCCAGCAGCCCGCCAGGATG	658
Db	1381	yGluProGlyProSerGlyProProGlyLysArgGlyPro-----	1394
QY	657	AGCAGTGCAGCTCCAGCGGCTGGGATCCGGCACAGCAGCCCTCTAGCCAGCGCGG	598
Db	1395	-----Pro-Gly-----ProAlaGlyP	1400
QY	597	CTTGGATGAGAAAGAGCTCAGCAGATGCCAAGACAGTCCCGCAGATGAAGGCCGG	538
Db	1400	to---GluGlyArgGlnGlyGluLysGlyAlaLysGlyGluAlaGlyLeuGluGlyPro-	1418
QY	537	CGCGGCCATAGCTCCAGCGCCAGTGTCTCACTGGCTG-----AGCTAGGAGCGGAC	485
Db	1419	-----ProGlyLysThrGlyProIleGlyProGlnGlyAlaProG	1432
QY	484	ACAGACCAGGCCCA---GCATGGACCAATGCCACGACCATGTCATCACTTCCTCTC	428
Db	1432	lyLysProGlyProAspGlyLeuArgGlyIleProGlyProValGly--GluGlnGlyL	1451
QY	427	TACCCCACTTCAGCAGCAGGAGCGGCACATAGGTATCCCTCGCGCCCAACACACCTC	368
Db	1451	euProGlySerProGly-----ProAspGlyProP	1461
QY	367	CAGGCGAAAGGTTAGCAGGTTGACAGCAAGAGCTGGGCTTCCTGGTCCGACAGCAGG-	309
Db	1461	roGlyPro-----MetGlyProProGlyLeuProGlyLeuLysGlyAspSerGlyProL	1479
QY	308	-----GGTCAAC	287
Db	1479	ysGlyGluLysGlyHisProGlyLeuIleGlyLeuIleGlyProProGlyGluGlnGly-	1498
QY	286	CATAGTGGG-----CCAGGCGGTAGGCTCAGGGGCGCTTCAGGCACCTCCA	239
Db	1499	GluLysGlyAspArgGlyLeuProGlyProGlnGlySerGlyProLysGlyGluGln	1518
QY	238	GAACGTCTCTCGGTCTCGGCTGTCTCCGAAGCTCGGCGCTCTCTCTCTCTCTCTCTCTCT	179





```
QY 2131 -----AGAGGAGCGCCAGCC-----CCAGCTGTGCAGCTAGCACCTCAG 2089
Db : : : : :
QY 573 roglyAlaLysGlyProGlyProGlyProGlyGluProAlaLeuSerGlyArg---LysG 592
QY 2088 CAGCAGAGGTGGCAGCAGAGAGCCACATTTACTTTGGCAGCAACAGAACTGGCGCCAG 2029
Db : : : : :
QY 592 lyAspGlnGlyProGlyProGlyProGlyPheProGlyProAlaGlyProA 612
QY 2028 CCGGAGCCCATGGGTATACAGAGAGCGGGAGCTGGGACCCAGTAGGCA----- 1976
Db : : : : :
QY 612 laGlyProPro---GlyTyrGlyProGlnGlyGluProGlyProLysGlyAlaGlnGly 631
QY 1975 -----GCCCTCCACCCCAATGTGCTGGAAGTTTCTACGCTCAGTATTG 1930
Db : : : : :
QY 631 alProGlyValLeuGlyProPro-----GlyGlu---AlaGlyLeuLysGly 646
QY 1929 GCCAAGTCGCTGTGTAATACTACCTCTGTAGCAAGTAATGGCAGCAGCCAGG 1870
Db : : : : :
QY 646 uProSerThrSer-----ThrProAspLeuG1 655
QY 1869 CTGCGGCAGACACCATATAGGCAGTCACAGACTGGCTGAGTGGACATGAGCCCAT 1810
Db : : : : :
QY 655 yProGlyProPro----- 660
QY 1809 AACAGGATGGGCGCCACCTGGGACAGCAGCA---AGGCACATTCAGGATGGCGAGTCC 1753
Db : : : : :
QY 661 -----GlyProProGlyGlnAlaGlyProArgGlyLeuProGlyLeuProGlyPr 677
QY 1752 A---GBCAGATGC-----CCGCGCCCGGAA----- 1731
Db : : : : :
QY 677 oValGlyLysCysAspProGlyLeuProGlyProAspGlyGluProGlyLeuProGluAl 697
QY 1730 -----CCACCTGGCTCGGTGGGTCTACCCACCACACAGCTAGCGAGACA 1684
Db : : : : :
QY 697 aglyCysProGlyProGlyProLys---GlyAsnGlnGlyPheProGlyThrLysGly 1716
QY 1683 TCACAGGACAGGCGCGGAGAGCGGGGTGGAGTGGGAGCAGGACCTGCTCCAGCA 1624
Db : : : : :
QY 717 Ser-----ProGlyCysProGlyGluMetGlyLysProGlyArgProGlyGlu 732
QY 1623 CCCACGTGTCATTAGGAGGAGGAGCTCCAGGCTTAGG----- 1585
Db : : : : :
QY 733 ProGlyLeuProGlyAlaLysGlyGluProSerValGlyArgProGlyLysProGlyLys 752
QY 1585 ----- 1585
QY 753 ProGlyPheProGlyGluArgGlyAsnAlaGlyLysAsnGlyAspIleGlyLeuProGly 772
QY 1584 -----CCTGGCAGGAAGCTGTGTATCATCAGGCTGCTCCTCAGCTGCTA 1546
Db : : : : :
QY 773 LeuProGlyLeuProGlyThrProGlyArg-----GlyGlyLeuAsp 786
QY 1545 GCACCTCAGTGTCCCTCGGTATTTGGCAGGAGACACCTGCTTCCCGGTGGTAGAG 1486
Db : : : : :
QY 787 GlyProProGlyAspProGly-----GlnProGlySerProGlyAlaLysG1 802
QY 1485 GAGGCCA---GTGTGTAGGCGAGGATCTGCAGGCTGAGAGGTGAACCCGGTAGGGCG 1429
Db : : : : :
QY 802 ySerProGlyArgCysIleProGlyProArgGlyThrGlnGlyLeuProGlyLeuAsnG1 822
QY 1428 GCTGAAGTGTTCACCGGCCACACTGTGGACAGGATGTGGACCGGACGAGG 1369
Db : : : : :
QY 822 yLeuLysGlyGlnProGlyArgGlyAspThrGly-----ProLysG1 837
QY 1368 AAAGCTGCCACATGGCCAAATAGACTCCTCGAGTGGCGGAATCGCTGCACAGCGGTCC 1309
Db : : : : :
QY 837 yAsp---ProGlyIleProGlyMet-----AspArgSerGlyValProGlyAspPr 853
QY 1308 ATGACACAGAGAAGACAGGAGATGCGGCACTGCAGGACACGCCCGCTGCCCATC 1249
Db : : : : :
QY 853 oGlyPro-----ProGly-----ThrProGlyCysProG1 863
QY 1248 GAAAGCCTTCATCATAGTGTCTCGGG-----CCTCGGTGCC 1210
```



Db 258 gGlyGlu-----AspGlyMet---AlaGlyAsnProGlySerValGlyPro---II 273  
QY 2129 AGGAGCGCCACGCCCGCCAGCTGTGCAG-----CTACGACCTCAGCAGCACA 2082  
Db 273 eGlyProProGlyProProGlyPheProGlySerSerGlyAlaLysGlyAspAlaGlyGI 293  
QY 2081 GCGTCGACGAGAGCA---CATTACTTTGGCAGCAGCAAGAACTGGCGCCAGCGG 2025  
Db 293 nSerGlyAspArgGlyProMetGlySerSerGlyAlaProGlyAsnAsnGlyAlaG 313  
QY 2024 GCAGCCCATGGGCGCTAACAG-----AG 2001  
Db 313 yGlnProGlyGlnProGlyArgProGlySerProGlyAspGlyAsnProGlyThrLy 333  
QY 2000 CGGGGAGCTGGGACCCAGTCAGCGCCCT----- 1970  
Db 333 sGlyGluMetGlyProAla---GlyGlyProGlySerProGlyPheProGlyProGlnGI 352  
QY 1969 -CCACCCCAATGTCTGGAAGTTTCTACGCTAGTATTGGCCAAAGTCGCTCTTGCA 1911  
Db 352 yProGlyAlaAlaGlyAla----- 359  
QY 1910 ATACTACCTGTAGCAAGTAATAATGGCGACCA---GACCCAGGCTGGGCGACAC 1855  
Db 360 -----GlnGlyAlaAlaGlyAsnLysGlyAspGlnGlyArgAspGlySerPr 375  
QY 1854 ATATAGCAGTGACAGACTGGCTGACATGGACAAATGGAGCCCATACAGGGATGGGCC 1795  
Db 375 oGlySerGlnGlyValGlyValGlyAlaAlaGlyGlnProGlyAlaGlnGlyGluArgGlyLe 395  
QY 1794 ACCTGGACAGCAGCAGGAGGACATCCAGGATGGCGAGGTCCA-----GCCAGATG 1744  
Db 395 uProGlyLeuProGlyGlnGlu-----GlyLysArgGlyProLeuGlyAlaGlyPr 413  
QY 1743 CCGCGGCGGAAACCCCTGG-GCCTGGTGGGCTCACCACACCC----- 1699  
Db 413 oProGlyProProGlyProSerGlyGluArgGlyLeuProGlyThrProGlyMetSerGI 433  
QY 1698 -----ACACG 1694  
Db 433 yArgProGlyAlaAlaGlyValAlaGlyLysAspGlyAlaArgGlyLeuGlnGlyGluAr 453  
QY 1693 TACGAGACATCAG-----GCAGAGGCGCCGACAGCGCGGGTGGAGTGGAG 1643  
Db 453 gGlyGluAlaGlyGlnGlyValAlaGlyAlaProGlySerAlaGlyGlnThrGlyAl 473  
QY 1642 CAGCGCACTGCTCCACGCCCTGCTCCATTAGGAAAGGAGCTCCAGGC---TTAGG 1586  
Db 473 aArg-----GlyAlaProGlyPheProGlyAlaAspGlyAlaProGlyGluArgGI 490  
QY 1585 GCCTGGCAGGAAGC---TGGTCATCAGGCTGTCTCTACTGCTAGCACCTCCAGTGTCC 1529  
Db 490 yPro-AlaGlyAsnProGlyValAspGlyAsnPro-----GlyGluGlnGlyProG 507  
QY 1528 TCGGTATTGGGCAACACCT-----GCTT 1502  
Db 507 lGlyIle---GlnGlyProProGlyLeuGlnGlyProSerGlyAsnLysGlyAspThrG 526  
QY 1501 CTCCCGGTGTAGAGGAGCCAGCTGTAGGCGAGGATCTGAGGCGCTCAGAGGTGAA 1442  
Db 526 lProSerGlyLysAspValGlyArgProGlyAlaValGly----- 541  
QY 1441 CCGGTGAGGCGGCTGAAGTGTGCACGAGGCGCACACTGTGGACAGGCATGTGGCAC 1382  
Db 542 -----ProArg-----GlyGluA 546  
QY 1381 GGCAGCCACAGGAAAGCTGCCACAC"GGCCAAATAGACTGCTGCTGCGATCCTG 1322  
Db 546 rgGlyProAlaGly-----GluAlaGlyLeuGluG 556  
QY 1321 CACCAGCGGTCCATGACCAGAGAGACAGGAGATGGCGCACTGCAGAGACACGCC 1262  
Db 556 lProProGly-----SerProGlyAlaGluGlyGluIleGlyAlaAla- 570

QY 1261 CAGGTCGCCCATCCGAGCGCTTTCATAGTGTCTCGGGCTCGTGCCTCCGCTCAGC 1202  
Db 571 --GlyArgProGlyGluGlnGlyPheGlnGlyLeuProGlyProAlaGlyPro----- 587  
QY 1201 TCTGGGCGCGCTCGTACAGCCCTCGCCCGCAGCAATCCGTGTAAACACGCGTGAAGGT 1142  
Db 588 -----ProGlyGluAlaGlyArgPro---GlyProValGly-----GluS 600  
QY 1141 CATGAGTCCCATCCAGCTGCACAGCTCAGCCAGCAGAGAGCGCGCGCTGCGGGCAT 1082  
Db 600 erGlyIleProGlyGluProGlyAlaSerGlyGluArgGlyGluArgGly-----A 617  
QY 1081 GCGGAGCAGCAGCTGTCAGCGGGGAGAGCGCGCCAGCTCCGAAAGCAAGCG 1022  
Db 617 laProGlyGluValGlyValSerGlySerArgGlyAlaProGlyGluArgGlyProSerG 637  
QY 1021 GCGCGGCGATGACAGCAGTGGGGCGCAGAGAGGGCGCGCACGCCCT----- 972  
Db 637 lyProSerGlyGluValGlyGlnAlaGlyProProGlyProAlaGlyAlaArgGlyAspA 657  
QY 971 -----CTGCTGGCTCGTGGGCGCCAGCGCTGCT-----C 941  
Db 657 laGlyAlaGlnGlyLeuIleGlyMetProGlyGluArgGlyProIleGlyArgAsnGlyP 677  
QY 940 CTCAGCCACAGCAGTGTGCTACGACAGGTGAGGAAGATGAGGGTGAGCGGCCAAA 881  
Db 677 roGln-----GlyAsnArgGlyLeuThrGlyGluA 687  
QY 880 GAGGCACTCTCTCTGGTGGTCCAGGTAGGGGCGCAGGCGCTGGTGTCCCATCAATGCC 821  
Db 687 rgGly---GlnAspGlyGluProGlyArg---ProGlyGlu----- 698  
QY 820 AGGAGGAGTAGCCAGCAGCGCCCAAGACTGATCATGAAGCATAGACAGACTAGGC 761  
Db 699 -----AlaGlyAlaProGlySerPro-----GlyGlnP 708  
QY 760 CTGGCGCACAGTGTCTGGTCCCGGAGAGAGGTTCAGAGA-----GCAGGGCTCCAGTGG 707  
Db 708 roGlyProSerGlyLeuValSerAlaLysAspArgGlyGluAlaGlyProAlaGlyG 728  
QY 706 AGTGAAGCACACT-----GCCACAGAGTCCAGCAGCGCCACGCCCGCATGAG 656  
Db 728 luProGlyProProGlySerSerGlyGlnArgGlyProAlaGly----- 742  
QY 655 CAGTGCAGCTCCAGGGCTGGGATCCGGGCGACAGCAGCGCTGTAGCAGCGCGCT 596  
Db 743 -----Ala-GlnGlyPro-----GlnGlyProThrGlyLeuSerGlyPro 755  
QY 595 TGG-----GATGAGAAGAGCTCAGCAGGAT 569  
Db 756 ThrGlyGluMetGlyGlnThrGlySerAspGlyLysAspGlyAlaLysGlyAspThrGly 775  
QY 568 GCCCAGGACAGTCCCGCATGAGGCGCGCGGCCCATAGCGTCCAGCGCCAGTGGTC 509  
Db 776 AlaArgGlyTyProGlyGluAlaGlyProIleGlyAlaProGly----- 790  
QY 508 ACTGGCTGAGCTTAGGAGGCGGACACAGCAGCAGCGCCAGCAGTCCAGTCCACCAATGCCACGAC 449  
Db 791 ---Asn-GluGlyArgGluGlyArgLysGlySerArgGlySerGlyIleProGlyAs 809  
QY 448 CATGGTCATGAATCTCTCTACCCCTTCCAGCAGCAGCGCGGCGGCATAGTGTAT 389  
Db 809 n-----SerGlyThrProGlyAspProGlyArgAlaGlyProProGlySe 824  
QY 388 GCGTGGCGCCAAACACACCTCCAGGCGCAAGGTTAGCA----- 351  
Db 824 rProGlyAlaGln-GlyProProGlyProSerGlyAlaThrGlyLeuSerGlyAspGlyG 844  
QY 350 -----GGTTGACACAGAGAGCTGGCTTTCCGGTGGC----- 318  
Db 844 lyGluArgGlyGluThrGlyProProGlyArgSerGlyGluProGlyAlaProGlyMeP 864

```

Db 330 rAspMetGlyValArgAlaGlyAlaLeuGlyLeuMetLeuAsnSerValValLeuGlyAl 350
QY 1294 CTTCTCTCTGTCATGGACCGCTGCTGCACGAGTTCGGCAGCAGCTATTTGGC 1353
Db 350 aThrSerLeuGlyValAspIleLeuAlaArgGlyValGly---GlyValLysArgLeuTr 369
QY 1354 CAGTGTGCAGCATTTCCTGTGGCTGCGGTGCGGCACATGCTGCTCCACACATGTGGCGT 1413
Db 369 pGlyIleValAsnPhelLeu---AlaIleCysLeuGlyLeuThrValLeuVa 386
QY 1414 GGTG-----1417
Db 386 lThrLysLeuAlaGlnHisSerArgGlnTyrAlaProGlyThrGlyGlyLeuGlnAspPr 406
QY 1418 -----ACAGCTTCAGCGCGCTCACCAGGTTTCACTTCACGCGCT 1458
Db 406 oLeuProProSerGlyGlyIleLysAlaGlyAlaLeuThrLeuPheserValLeuGlyI 426
QY 1459 GCAGATC-----CTGCCCTACACACTGGCTTCCTTACCACCGGGAGAA 1503
Db 426 eProLeuAlaIleThrTyrSerIleProPheAlaLeuAlaSerIlePhe-----442
QY 1504 GCAGGTGTCTCTGCCCAATACCAGGGGACACTGGAGGTGCTAGCAGTGCAGCAGCCT 1563
Db 443 -----SerSerThrSerGlyAlaGlyGlnGlyLeuSerLe 454
QY 1564 G-----ATGACCAGCTTCCTCCGACGCGCC 1587
Db 454 uGlyValLeuAsnLeuAlaIleValIleProGlnMetPheValSerValLeuSerGlyPr 474
QY 1588 TAAGCCTCGA-----GCTCCCTTCCTTAATGGACACGTGGGTGC 1626
Db 474 oTrpAspAlaLeuPheGlyGlyAsnLeuProAlaPheValValGlyAlaValAlaAl 494
QY 1627 TGGAGGCGAGTGGC-----CTGCTCCACCTCCACCC 1657
Db 494 aLeuAlaSerGlyIleLeuSerMetIleLeuLeuProSerProPro 509

RESULT 33
097405 PRELIMINARY; PRT; 1378 AA.
AC 097405;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Collagen pro alpha-chain precursor.
GN HDCOL1 ALPHA.
OS Haliotis discus (Abalone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
OC Haliotidae; Haliotis.
OX NCBI_TaxID=36094;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RX MEDLINE=99234051; PubMed=10215888;
RA Yoneda C., Hirayama Y., Nakaya M., Matsubara Y., Irie S., Hatae K.,
RA Watabe S.;
RT "The occurrence of two types of collagen proalpha-chain in the abalone
RL Haliotis discus muscle.";
RL Eur. J. Biochem. 261:714-721(1999).
DR EMBL; AB017600; BAA75668.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; COLFI; 1.
DR Pfam; PF01391; Collagen; 17.
DR ProDom; PD000007; Collagen; 1.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; WVC; 1.
DR KW Signal.
FT SIGNAL.
FT SIGNAL.
SQ SEQUENCE 1 18 POTENTIAL.
SEQUENCE 1378 AA; 132583 MW; 6EEE34A65B52E634 CRC64;

```

```

Alignment Scores:
Pred. No.: 2,56e-13 Length: 1378
Score: 333.00 Matches: 299
Percent Similarity: 30.71% Conservative: 60
Best Local Similarity: 25.58% Mismatches: 389
Query Match: 5.35% Indels: 423
DB: 5 Gaps: 66

US-09-759-143-110 (1-3410) x 097405 (1-1378)
QY 3077 TGTGCTTCTGGTCTCAGTAGCTCCAAACAGGGTTGTGGAGCTGTGGGAAAGTTGGG 3018
Db 84 CysAlaSerCysProGluLysAlaSerAsn-----ThrGluGlnAlaGlnAsnThr---100
QY 3017 GGTAGGGAAAGTGGGGGTAGGGGAAATTTTGGCAGTGCCTTCATCAGCCACGCTCCTA 2958
Db 101 ---GlnGlnValLysAlaAlaProAlaTyrSerProGlnTyrTyrSerProGlyAla 119
QY 2957 GAGAGAGTAGAGGGAGTGAAGTGGGGGAACAGGCTGGGGCCCAAGAGAA---GAGGGG 2901
Db 120 ProGlyProArgGlyProProGlyProGlyGlnProGlyProGlyProGlnGlyPheGlnGly 139
QY 2900 TGGTTAGGAAGCCCTTGAGACCTGAAGCCGCCACCTCTACCTTCTTCAACACCCCTAAC 2841
Db 140 MetArgGlyGluSerGlyAspProGlyProGlySerThr-Gly-----154
QY 2840 CTTGGGTACACGATTTGGAATATCATTTGGGATGAGTAGAATTTCCAAGTCTCTGGGT 2781
Db 155 -----G1 155
QY 2780 TAGGCATTTTGGGGGCCAGACCCAGGAGAGAAGATTTCTGGCAATGATCAGCCCAATG 2721
Db 155 nargGlyPheAlaGlyProProGlySer-----164
QY 2720 ACCAGCTATCTCAGGGACCTG---ATTGTTGGGATCCCCACCTCAGCCACATCTGAT 2664
Db 165 -ProGlyLeuAspGlyGluGluGlyMetProGlyAspProGlyProAlaGlyProLeuG1 184
QY 2663 ACACCAACACAGAAAAGCTAGCAATGCTTCTACTTTGTTTAAATAAATAGTTAA 2604
Db 184 ySerProGlySerSerGlyArgProGlyMetPro-----195
QY 2603 ATATTAAATGCTGTGCTCTGTGATGGCAACAGAGGACCAACAGCCACATCTGAT 2544
Db 196 -----GlyMetProGlyProGlyHisArgGlyPr 206
QY 2543 AAAAGGTAAAGGGGGTGGATCACAAAAGACAGTGTGTGGCTCAGGGGACCTGGT 2484
Db 206 oSerGlyArgProGlyArgAspGlyGluAlaGlyLysGln-----GlyAsp-----221
QY 2483 TCTTGTGTGTTGCCCTCAGGACTCTCCCTTACAAATAAGTCATATGTTCAAAATCCCAT 2424
Db 221 -----221
QY 2423 GGAGAGGTGTTTCATCTTAGAACTCCCATGCAAGAGCTACATTAAACGAAGTGCAGGT 2364
Db 221 -----221
QY 2363 TAAGGGGCTT-----AGAGATGGGAACACAGGTGACTGATTTATTTCAGCTCCCAAAA 2310
Db 222 -LysGlyIleProGlyAlaAlaGlyProProGlySer-----SerGlyProGlnG1 238
QY 2309 CCCTTCTCTAGTGTGTCTCAACTAGGAGGCTAGTGTGTTAAACCTTGAGCCCTGGGTATPCC 2250
Db 238 yProArg-----GlySerPr 243
QY 2249 ACCTCCAGATCCCGGCATTCCAGTGCATGAGGCCCTTCTGGCTCTCTGATATAGTCCA 2190
Db 243 oGlyGluArgGlyArgAspGlySerProGlySerProGly-----IleAr 258
QY 2189 GACTGAAACCCCTTGGGAAGCCTCCAGTCCAGGAGCCCTAGAGACTGGGGAGAGAGGAG 2130

```

```

Db      351 GlyValGluPheLeuGlyLysLysIleGly---GlyAlaLysArgLeuTrpGlyIleLeu 369
Qy      1364 GCTTCCCTGGGCGCGGTGCCACATGCTGTCCACAGTGTGGCGCGTGGT----- 1417
Db      370 AsnPheValLeu-----AlaIleCysLeuAlaMetThrIleLeuValThrLysMet 386
Qy      1418 -----ACAGCTTCAGCGCCCTACCGGGTTCACC----- 1447
Db      387 AlaGluLysSerArgGlnHisAspProAlaGlyThrLeuMetGlyProThrProGlyVal 406
Qy      1448 -----TTCTCAGCCCTCCAGATC----- 1465
Db      407 LysIleGlyAlaLeuLeuPheAlaLeuGlyIleProLeuAlaAlaThrPheSer 426
Qy      1466 CTGCGCTACACACTGGGCTCCCTCTACACCGGGAGAG----- 1504
Db      427 IleProPheAlaLeuAlaSerIlePheSerSerAsnArgGlySerGlyGlnGlyLeuSer 446
Qy      1505 -----CAGGTGTTCTCTGCCCAATACCGAGGGGAGTGCCT 1546
Db      447 LeuGlyValLeuAsnLeuAlaIleValProGlnMetLeuValSerLeuValGlyGly 466
Qy      1547 AGCAGTAGGACACCTGATGACAGCTTCTGCGAGCGCCCTAAGCCTGGAGTCCCTTC 1606
Db      467 ProTrpAspLeuPheGlyGlyGlyAsnLeuProGly-----Phe 480
Qy      1607 CCTAATGACACGTGGTGTGGAGGAGTGC-----CTGCTCCACCT 1651
Db      481 ValValGlyAlaValAlaAlaAlaSerAlaValLeuAlaLeuThrMetLeuProSer 500
Qy      1652 CCACCGCGCTCTCGGGGCTCTGCTGTGATGCTCTCCGTACGTGTGGTGGGT 1708
Db      501 ProProAla-----AspAlaLysProAlaValAlaMetGly 512

RESULT 32
Q9XHL6 PRELIMINARY; PRT; 524 AA.
AC Q9XHL6;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Sucrose transport protein SUT1.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosida I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99291546; PubMed=10363367;
RA Tegeder M., Wang X.D., Frommer W.B., Offler C.E., Patrick J.W.;
RT "Sucrose transport into developing seeds of Pisum sativum L.";
RL Plant J. 18:151-161(1999).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF109922; AAD1024.1; -.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar.tr.1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
SQ Transmembrane.
KW SEQUENCE 524 AA; 55279 MW; B5CE3F880D2C1E6B CRC64;

```

## Alignment Scores:

```

Pred. No.: 2.18e-13 Length: 524
Score: 333.00 Matches: 140
Percent Similarity: 39.3% Conservative: 79
Best Local Similarity: 25.18% Mismatches: 200
Query Match: 5.19% Indels: 137
DB: 10 Gaps: 20

```

US-09-759-143-110 (1-3410) x Q9XHL6 (1-524)

Qy 256 CCCCTGAGCCCTACCGCGCTGCGCCCATATGCTCCAGAGGCTGTGGGTGAGCGCCCTGCT 315

```

Db      3 ProLeuSerSerThrLysGlnIleAsnAsnAsnAsnLeuAlaLysProSer 22
Qy      316 GCGGCACCGGAAGCCAG-----CTCTTCTGCTGCA 348
Db      23 Leu-HisValGluThrGlnProLeuGluProSerProLeuArgLysIleMetValAla 42
Qy      349 CTGTCTAACCTTTGGCTGGAGGTGTGTGGCCGACGATCATCATATGTGCCGCTCT 408
Db      42 aserIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeuThrPro 62
Qy      409 GCTGTGGAGTGGGGGTAGAGAGAGTTCATGACCATGTGCTGGGCGATTTGGTCCAGT 468
Db      62 rValGlnLeuLeuGlyIleHisThrTrpAlaAlaIleTrpLeuCysGlyProI 82
Qy      469 GCTGGCGTGGTGTCTGCTCCGCTAGGCTCAGCAGTGCAGCAGTGCCTGGAGCGCTA 528
Db      82 eserGlyMetLeuValGlnProValValGlyTrpHisSerAspArgCysThrSerArg 102
Qy      529 TGGCCCGCGCGCCCTTCATCTGGGCTGCTCTTGGGCTGCTGCTGAGGCTCTTTCT 588
Db      102 eGlyArgArgProPheIleAlaAlaGlySerIleAlaValAlaIleAlaValPhe 122
Qy      589 CATCCCAAGCGCGCTGCTAGCAGGCTG-----CTGTGCC 627
Db      122 uile-----GlyTrpAlaAlaAspLeuGlyHisSerPheGlyAspAsnLeuAsp 139
Qy      628 GGATCCACGCGCTGGAGCTGCTCATCTGCTGGCGTGGGCTGCTGCTGCTGCTG 687
Db      139 sLysValArgProArgAlaIleGlyIlePheValValGlyPheTrpIleLeuAspVal 159
Qy      688 TGGCCAGGTGTGCTTCACTCCACTGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTG 744
Db      159 aAsnMetLeuGlnGlyProCysArgAlaLeuLeuGlyAspLeuCysAlaGlyAsn 179
Qy      745 GGACCACTGTGCGCAGCGCTACTCTGTCTATGCTTTCATGCTTTCATGCTTGGG 804
Db      179 nArgLysThrArgAsnAlaAlaPhePheSerPheMetAlaValAlaGlyAsnVal 199
Qy      805 GGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db      199 uGlyTrpAlaAlaGlyAlaTrpSerLysLeuTrpHisValPheProPheTrpLys 219
Qy      841 TGGCTGCGCGCTTACCTGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db      219 uAlaCysAsnValTrpCysAlaAsnLeuLysSerCysPhePheLeuSerIleAla 239
Qy      901 CTCTCTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942
Db      239 uThrValLeuAlaThrAlaAlaLeuIleTrpValLysGluThrProLeuIleAla 259
Qy      943 GGCAGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1002
Db      259 sAlaValAla-----ThrAlaGluAspGlySerAsnGlyGlyMet----- 273
Qy      1003 CTGCTGTCCATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
Db      274 -----ProCysPheGlyGlnLeuSerGlyAlaPheLysGluLeu----- 286
Qy      1057 CCGGCTGCACACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116
Db      287 -----LysArgProMetTrpIleLeuLeuValTh 297
Qy      1117 GCTGTGAGCTGGATGGCACTCATGACCTTACAGCTTTTACACGAGATTTCGTGGCG 1176
Db      297 rCysLeuAsnTrpIleAlaTrpPheProPheLeuPheAspThrAspTrpMetGly 317
Qy      1177 GGGCTGTACAGGCGCTGCCAGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1233
Db      317 sGluValTrp-----GlyGlyThrValGlyGluGlyHisAla 330
Qy      1234 TGATGAAGGCGCTCGGATGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1293

```

```

QY 1475 ACACCTGGCTCCTACACCGGAGAGAGAGGTGTTCTTCCGCCCAATACCGAGGGAC 1534
Db      |||:::  ::::
QY 384 -----GlnHisGlnArgGlnHis----- 389
QY 1535 ACTGGAGGTGTAGCAGTACAGGACAGCCTGATGATGACACAGCTTCTCCAGGCGCCCTAAAGCCT 1594
Db      |||:::  ::::
QY 1595 GGAGCTCCCTTCCCTTAATGGACACGTGGTGTGGAGCAGTGGCGCTCTCCACCTCCCA 1654
Db      |||:::  ::::
QY 1655 CCGCGCTCTCGGGGGCCTCTCGCTGTGATCTCTCCGTACCTGGTGGTGGGAGGCC 1714
Db      |||:::  ::::
QY 399 SerAla-----GlyValAlaGlyAlaLeuSerLeuPheSerIleLeuGlyIlePro 416
QY 1715 ACCGAGGCCAGGGT-----GTTCCG-----GGC 1738
Db      |||:::  ::::
QY 417 LeuSerIleThrPheSerIleProPheAlaLeuAlaSerIleTyrSerSerGlySerGly 436
QY 1739 CGGGCATCTCGCTGACCTCGCATCTGATAGTGCCTTCCGCTCTCCAGGTGCC 1798
Db      |||:::  ::::
QY 437 AlaGlyGlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValProGlnMetIle 456
QY 1799 CATTCCCTGTTTANGGGCTCCATCTGCTCCAGCTC-----AGCCAGTCTGTCACTGCCCTAT 1852
Db      |||:::  ::::
QY 457 ValSerValLeuAlaGlyProPheAspSerLeuPheGlyGlyGlyAsnLeuProAlaPhe 476
QY 1853 ATGGTG-----TCTGCCGAGGCGCTGGTCTGGTGGCCATT 1888
Db      |||:::  ::::
QY 477 ValValGlyAlaIleSerAlaAlaIleSerGlyValLeuAlaIle 491

```

## RESULT 31

```

Q43653 PRELIMINARY; PRT; 516 AA.
AC Q43653;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Sucrose transport protein.
OS SUT1.
GN Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanaceae; Solanum.
OX NCBI_taxid=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DESIREE; TISSUE=LEAF;
RX MEDLINE=94146554; PubMed=8312741.
RA Riesmeyer J., Willmitzer L., Frommer W.B.;
RT "Potato sucrose transporter expression in minor veins indicates a role
RT in phloem loading.";
RL Plant Cell 5:1591-1598(1993).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: X69165; CAA48915.1;
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR TIGRFAMs: TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 516 AA; 54831 MW; 4FD06C095E49A377 CRC64;

```

## Alignment Scores:

```

Pred. No.: 2.18e-13 Length: 516
Score: 333.00 Matches: 133
Percent Similarity: 41.93% Conservative: 93
Best Local Similarity: 24.68% Mismatches: 201
Query Match: 5.19% Indels: 112
DB: 10 Gaps: 21

```

us-09-759-143-110 (1-3410) x Q43653 (1-516)

```

QY 278 CCACCTATGTCTCAGAGGCTGTGGTGAGCCGCTCTGCTGGGACACCGGAAAGCCAGCTC 337
Db      |||:::  |||
QY 24 ProLeuAlaProSerLysLeuTrp-----LysIle 33
QY 338 TTGCTGTGAACCTGCTTAACCTTTTGGCTGAGGTGTGTTTGGCCGAGGCATCACCTAT 397
Db      |||:::  |||
QY 34 IleValValAlaSerIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeu 53
QY 398 GTCCCGCTCTGTCTGGAGTGGGGTAGAGGAGAGTTCATGACCATGTGTCTGGGC 457
Db      |||:::  |||
QY 54 LeuThrProTyrValGlnLeuLeuGlyIleProHisLysPheAlaSerPheIleTrpLeu 73
QY 458 ATTGGTCCAGTGTGGGCTGTGTCTGCTCCGCTCTAGGCTCACCAGTACACCATGG 517
Db      |||:::  |||
QY 74 CysGlyProIleSerGlyMetIleValGlnProValValGlyTyrTyrSerAspAsnCys 93
QY 518 CGTGGACGTATGGCCCGCCGCTTTCATCTGGGCACCTGTCTGGGCATCTCTGCTG 577
Db      |||:::  |||
QY 94 SerSerArgPheGlyArgArgProPheIleAlaAlaGlyAlaAlaLeuValMetIle 113
QY 578 AGCCTCTTCTCATCCCAAGGCGGCTGTGTAGAGGGCTG-----GlyPheAlaAlaAspLeuGlyHisAlaSerGlyAsp 130
Db      |||:::  |||
QY 620 ---CTGTGCGCGGATCCAGGCGCTGTGGAGCTGTGTGCTCATCTCTGCTGGGCTG 676
Db      |||:::  |||
QY 131 ThrLeuGlyLysGlyPheLysProArgAlaIleAlaValPheValValGlyPheTrpIle 150
QY 677 CTGGACTTCTGTGGCGAGGTGTCTACTTCCAGTGGAGGCGCTGTCTCTGCTGCTCTTC 736
Db      |||:::  |||
QY 151 LeuAspValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuAlaAspLeuSer 170
QY 737 CGSGACCCCGGAC---CACTGTGCGCAGGCTCTCTGTCTATGCTCTATGCTATGATCAGTCTT 793
Db      |||:::  |||
QY 171 GlyGlyLysSerGlyArgMetArgThrAlaAsnAlaPhePheSerPheMetAlaVal 190
QY 794 GGGGCTGTCTGGCTACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832
Db      |||:::  |||
QY 191 GlyAsnIleLeuGlyTyrAlaAlaGlySerTyrSerHisLeuPheLysValPheProPhe 210
QY 833 ---GACACCACTGCGCTGCGCCCTACCTGGCAGGAGGAGGAGTCTCTTT---GGC 886
Db      |||:::  |||
QY 211 SerLysThrLysAlaCysAspMetTyrCysAlaAsnLeuLysSerCysPhePheIleAla 230
QY 887 CTGCTCACCTCATCTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943
Db      |||:::  |||
QY 231 IlePheLeuLeuLeuSerLeuThrIleAlaLeuThrLeuValArgGlnAsnGlnLeu 250
QY 944 GCAGCGCTGGGCGCCACCGAGCAGCAGGCGGTGTGGGCGGCTGCTGCTGCTGCTGCTGCTG 1003
Db      |||:::  |||
QY 251 ProGluLysAspGluGlnGluIleAspGluLysLeuAlaGlyAlaGly----- 266
QY 1004 TGTGTCTCATGCGGCGCGCTGTGCTTTCGGAACCTGGGCGGCTGCTCTCTCTCTCTCTG 1063
Db      |||:::  |||
QY 267 -----LysSerLysValProPhe-----PheGlyGluIlePheGlyAlaLeu 280
QY 1064 CACCACTGTGTGCGCGCATGCGCGCACCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTG 1123
Db      |||:::  |||
QY 281 LysGluLeu-----ProArgProMetTyrPheLeuLeuValThrCysLeu 296
QY 1124 AGCTGGATGGCACTCATGACCTTTCAGCTGTGTTTACAGGATTTCTGGGCGGAGGGCTG 1183
Db      |||:::  |||
QY 297 AsnTrpIleAlaTrpPheProPheLeuTyrAspTrpPheMetAlaLysGluVal 316
QY 1184 TACCAGGCGGTGCCAGAGCTGAGCGGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1243
Db      |||:::  |||
QY 317 PheGlyGly-----GlnValGlyAsp-----AlaArgLeuTyrAspLeuGly 330
QY 1244 GTTCGATGGGCGAGCTGGGGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1303
Db      |||:::  |||
QY 331 ValArgAlaGlyAlaMetGlyLeuLeuGlnSerValValLeuGlyPheMetSerLeu 350
QY 1304 GTCATGGACCGGCTGTGTGCGAGGATTCGGCACTCCAGCAGTCTATTTGGCAGTGTGGCA 1363

```

US-09-759-143-110 (1-3410) x Q9ZTB9 (1-512)

QY	332	CAGCTCTGCTGGTGAACCTGCTAACCTTTGGCTGCAGGTGTTTGGCCGAGCATC	391
Db	33	LysLeuIleLeuValAlaAlaIleAlaAlaGlnPheGlyTrpAlaLeuGlnLeu	52
QY	392	ACCTATGTGCGCCTCTGCTGCTGAATGGGGGTACAGCAGAAAGTTCATGACCATTGGT	451

US-09-759-143-110 (1-3410) x 09ZTB9 (1-512)

QY	332	CAGCTCTGCTGGTCAACCTCTAACTTTGGCGCTGGAGGTGTGTTGGCCGAGGCATC	391
Db	33	LysLeuIleLeuValAlaAlaIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu	52
QY	392	ACCTATGTGCGCGCTCTGCTGCTGGAAGTGCGGGGTAGAGGAGGTTCATGACCATGGTG	451



QY 2159 AGCGAGCCCTAGAGACTGGGGAGAGAGG-----AGAGGGAC 2124  
Db 657 -GlyLeuProGlyAspGlyAspGlyGluValGlyProArgGlyLe 676  
QY 2123 GCCCAGCCCGCAGCTGTGCAG-----CTAGCCACCTCAGCAGCAGGTCGAGCAG 2070  
Db 676 uProGlyLeuProGlyProArgGlyLeuLeuGlyProLysGlyProProGlyProGly 696  
QY 2069 AGAGCCACATTATTGGCAGCAACAACTGGCGGCAGCCCGCAGCCCATGGGC 2010  
Db 696 yProPro-----Gly-ValThrGlyMetAspGlyGlnProGly---ProLysGlyA 712  
QY 2009 TAACAGGA---CCGGGGAGCTGGGACCCAGTAGCAGCAGCCCTCCACCCCAATGTCTGG 1953  
Db 712 snValGlyProGlnGlyLeuProGlyProProGlyGlnGlnGlyAsnProGly----- 729  
QY 1952 AAGTTTTCTACCTCAGTATTGGCAAGTCGCTCTGTCAATACTACCTGTGTAGCAA 1893  
Db 730 -----AlaGlnGlyLeu-ProGlyProGlnGlyAlaIleGlyProProGlyGlu 745  
QY 1892 AGTAATGGCAGCAGACCCAGGCTGGCGGCAGACACCATATAGCAGTGCAGACTGCG 1833  
Db 746 LysGlyProLeuGlyLysProGlyLeuProGlyMetProGlyAlaAspGlyProProGly 765  
QY 1832 TGAGCTGCACATGAGGCCATTAACAGGATGGCGCCACCTGGG---ACAGCAGAGAG 1776  
Db 766 HisProGlyLys-----GluGlyProProGlyGluLysGlyGln 779  
QY 1775 CACTATCAGGATGGCGAGGTCAGGATGCCCGCCCGCCGGAA----- 1731  
Db 780 GlyProProGlyProGlnGlyProIleGlyProGlyProArgGlyValLysGlyAla 799  
QY 1731 ----- 1731  
Db 800 AspGlyIleArgGlyLeuLysGlyThrLysGlyGluLysGlyGluAspGlyPheProGly 819  
QY 1730 -----CCACCTGGCTGG 1716  
Db 820 PheLysGlyAspMetGlyIleLysGlyAspArgGlyGluIleGlyProGlyProArg 839  
QY 1715 TGGGCTCACCACCCACAGCTACGGAGACATCACAGCAGAGCCCGCAGAGCGCGG 1656  
Db 840 --GlyGlu-----AspGlyProGlyLysGlyArgGly 851  
QY 1655 GTGGAGTGGAGCAGGCACTGCTCCAGCACCACCTGTCATAGGGAAG---GGAG 1599  
Db 851 yGlyProAsnGlyAspProGlyProLeuGlyProThrGlyLysGlyLysLeuGlyV 871  
QY 1598 CTCAGGCTTAGGG-----CCTGGCAGGAAGCTGCTCATCAGGCTGCTCCTCAGTCTAG 1545  
Db 871 alProGlyLeuProGlyTyProGlyArgGln-----G 882  
QY 1544 CACTTCCAGTGTCCCTCGGTATT-----TGGCAGGAACACCTGCTTCTCCCGGTGT 1491  
Db 882 lyProLysGlySer-IleGlyPheProGlyPheProGlyAlaAsnGlyGluLysGlyGly 901  
QY 1490 AGAGGAGCCAGCTGTAGGCGCAGGACTCTCGAGCGGTGAGAGTGTAACCCGCTGAGG 1431  
Db 902 ArgGlyThrPro-----GlyLysProGly----- 909  
QY 1430 CGGCTGAAGCTGTACCACGCCACACTGTGGGACAGGCATGTGCAACCGCAGCCACAG 1371  
Db 910 -----ProArg 911  
QY 1370 GGAAGCTGCCACACTGCCCAATAGACTGCTCGAGTCCGGAATCGCTGCACCCAGCCGCT 1311  
Db 912 GlyGln-----ArgGlyProThrGly 918  
QY 1310 CCATGACCAGAGAGA-----AGACAGGG---AGATGGCGCACT 1275  
Db 919 ProArgGlyGluArgGlyProArgGlyIleThrGlyLysProGlyProLysGlyAsnSer 938  
QY 1274 CGAGAACACGCCCA---GGTGCTCCCATCCGAAGCCCTTCATCATGTCTCCGGGCT 1218

Db 939 GlyGlyAspGlyProAlaGlyProAlaGlyProGlyGluArg-----GlyPro 952  
QY 1217 CGGTCCCGGCTCAGCTGTGGCAGCCCTGTGTACAGCCCTCGCCCAAGAAATCCGTGT 1158  
Db 953 AsnGlyProGlnGlyProThrGlyPheProGlyProLysGlyProProGlyProGly 972  
QY 1157 AAACAGCGCTGAAGTCTAGTGCAC-----LeuProGlyHisProGlyGlnArgGlyGluThrGlyPhe 988  
Db 973 LysAspGly-----LeuProGlyHisProGlyGlnArgGlyGluThrGlyPhe 988  
QY 1130 TCCAGCTGCACAGCTCAGCCAGAGAGCCGCGCAGGCTGCGGGCATGCGGCAGACACA 1071  
Db 989 GlnGlyLysThrGlyProProGlyProGlyVal---ValGlyProGlnGlyProThr 1007  
QY 1070 -----GCTGTGTCAGCCGCGGAGCAGCCCGCCAGGTTCCGGAAGCAGCAGCGG 1020  
Db 1008 GlyGluThrGlyProMetGlyGluArgGlyHisProGlyProGlyProGlyProGlyGlu 1027  
QY 1019 -----CCGCGCATGGACAGCAGTGGGCGCAGAGGAGGCGGCGCAGCCCTTCGTG 969  
Db 1028 GlnGlyLeuProGlyValAlaGlyLysGluGlyThrLysGlyAspPro----- 1043  
QY 968 CTGGCTCGGTGGGCGCCAGCG---CTGCTCTCTCAG-----CCACAGCAGTGTGGCTG 918  
Db 1044 -----GlyProAlaGlyLeuProGlyLysAspGlyProPro----- 1055  
QY 917 CTACCCAGGTGAGAGATGAGGCTGAGCAGCCAGCAGCCCAAGAGGCGCAGT----- 873  
Db 1056 -----GlyLeuArgGlyPheProGlyAspArgGlyLeuProGlyProValGly 1071  
QY 872 -----CCTCTGGGTGCCAGGT-----AGGGGCCA 846  
Db 1072 AlaLeuGlyLeuLysGlySerGluGlyProGlyProGlyProGlyProAlaGlySerPro 1091  
QY 845 GGGCACTGTGTCCAGCTCAATGGCAGGAGGAGTAGCCAGCAGGAGCCCGCAGACTGA 786  
Db 1092 GlyGluArgGlyProAla-----GlyAlaAlaGlyProIleGlyLeuPro----- 1106  
QY 785 TCATCAAGGCATAGACAGATAGGCTGTGGCAGCTGTGTCGGGCTCCCGAAGAGGTCAG 726  
Db 1107 -----GlyArgProGlyProGlnGlyPro----- 1114  
QY 725 AGAGAGGCGCTTCCAGTGGAGTGAAGCAGACCTGTCACAGAGAGTCCAGCAGCCACGC 666  
Db 1115 ---ProGlyProAlaGlyGluLysGlyValProGlyGluLysGlyProGlnGlyPro-Al 1133  
QY 665 CCAGATCAGCAGTCCAGCTCCAGGCGCTGGGATCCGGGCGCAGCAGCAGCCCTGTAGCC 606  
Db 1133 aclyArgAsp-----GlyLeuGlnGlyPro---ValGlyLeuProGlyProAlaGlyPr 1150  
QY 605 AGCCGCGCTTGGGATGAG-----AAAGAGGCTCAGCAGGA 570  
Db 1150 oValGlyProProGlyGluAspGlyAspLysGlyGluIleGlyGluProGlyGlnLysGly 1170  
QY 569 TGCCCAAGCAGAGTCCCGCAGATGAAGGCGCGCGGCCCATAGCTCCACGCCAGTGTGT 510  
Db 1170 ySerLysGlyAspLysGlyGluGlnGlyProProGlyProGlyProGlyProGly 1189  
QY 509 CACTGGCTGAGCCT---AGGAGCGGCGCAGACAGCAGCCCGCAGCAGTGGACCAATGCCCA 453  
Db 1189 roIleGlyGlnProGlyProSerGlyAlaAspGlyGluPro-----GlyProArgGlyG 1207  
QY 452 GCACCATGCTCATGAATCTCTCTCTACCCCTCTCCAGCAGCAGAGCGCGCAGCATAGG 393  
Db 1207 lGlnGlyLeu-----PheGlyGlnLysGlyAspGlyGluGlySerArgGly----- 1221  
QY 392 TGATGCTCGCGCCCAACACACCTCCAGGC---CAAGGTTAGCAGGTTGACCAACAAGA 336  
Db 1222 --PheProGly-ProProGlyProValGlyLeuGlnGlyLeuProGlyProProGlyGlu 1240  
QY 335 GCTGGCTTTCGGGT-----GCCCA 315

```

QY 338 ----AGAGCTGGCTTTCCGGTCCCGCAGCAGCGCGCTCACCCACAGCCTCTGGACCATAG 282
Db 1134 GlyArgAspGlyLeuGlnGlyProVal-----GlyLeuProGlyProAlaGlyPro-Va 1151
QY 281 TGGG---CAAGCGGGTAGGGCTCAGAGGGCGGTTCAGGCACTCCA---GAAGCTGCTCG 228
Db 1151 rGlyProGlyGluAspGlyAspGlyGlyGluGlyGluProGlyGlnLysGlySe 1171
QY 227 TCTCGCTCTGCTCAGAGCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 168
Db 1171 rLysGlyAspLysGlyGlnGlnGlyProGlyProGlyProGlyProGlyProGlyProGly 1191
QY 167 TCAGCAGCGCGCCCTTTCTGCG----- 145
Db 1191 yGlnProGly-ProSerGlyAlaAspGlyGluProGlyProGlyProGlyProGlyLeuP 1211
QY 144 -----AGCCTTTGCTGGCGGTCCAGC 123
Db 1211 heGlyGlnLysGlyAspGlySerArgGlyPheProGlyProGlyProGlyProValGlyL 1231
QY 122 TTCTCAGCCATGCTCAACACCTGCTGCTGCTGGGCGACCTCAGTGGGCGACAGTCTCATC 63
Db 1231 euGlnGlyLeuProGlyProGlyProGlyGluLysGlyGlu---ThrGlyAspValGlyGlnM 1250
QY 62 ACTCAGATCCTGCG-----CGAGGCGCGCGCTGTCAACCGGAGCC 22
Db 1250 etGlyProGlyProGlyProGlyProGlyProGlyProGlyProGlyProGlyAla 1267

RESULT 29
Q60467 PRELIMINARY; PRT; 1840 AA.
ID AC Q60467;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pro-alpha-1 type V collagen.
OS Crictulus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10030;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92105142; PubMed=172213;
RA Greenspan D.S., Cheng W., Hoffman G.G.;
RT "The pro-alpha-1(V) collagen chain: Complete primary structure,
RT distribution of expression, and comparison with the pro-alpha-1(XI)
RT collagen chain.";
RL J. Biol. Chem. 266:24727-24733(1991).
DR EMBL; M76730; AAA37002.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib.collagen.C.
DR InterPro; IPR001791; Laminin.G.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 18.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 2.
DR ProDom; PD002078; Fib.collagen.C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW Collagen.
SQ SEQUENCE 1840 AA; 184174 MW; 32C56821EF64CE75 CRC64;

```

## Alignment Scores:

```

Pred. No.: 2,15e-13 Length: 1840
Score: 334.50 Matches: 302
Percent Similarity: 29.66% Conservative: 59
Best Local Similarity: 24.82% Mismatches: 392

```

```

Query Match: 5.37% Indels: 464
DB: 11 Gaps: 68
US-09-759-143-110 (1-3410) x Q60467 (1-1840)
QY 3068 GGTCTCTGAGTAGTCCAAACAGAGGTTGTGGAGCTGTGGGGAAGTTGGGGTAGGGGA 3009
Db 433 GlyProGlyMetProAlaAsnGlnAspThrIleTyrGluGlyIleGlyGlyProArgGly 452
QY 3008 AAGTTGGG---GTAGGGGAAATTTTGGGCGAGTCCCTTCATCAGCCAGTCCCTAGACAGA 2952
Db 453 GluLysGlyGlnLysGlyGlu-----ProAlaIleIleGluProGlyMet---Leu 468
QY 2951 GTAGAGGG----- 2940
Db 469 IleGluGlyProGlyProGlyProGlyProAlaGlyLeuProGlyProGlyThrThr 488
QY 2939 GGAAGTGGGGGAACACAGGCTGGGCCAAGAGAGAGGGGTGTAGGGAAGCCCTTACAGA 2880
Db 489 GlyProThrGlyGlnValGlyAspProGlyGluArgGlyProProGlyArgProGlyLeu 508
QY 2879 CCTCAAGCC-----CCACCCTCTACCTTCCCTTCAACACCCTAACCTTGGG 2835
Db 509 ProGlyAlaAspGlyLeuProGlyProGlyProGlyThrMetLeuMetLeuPro----- 525
QY 2834 TAACAGCATTTTGAATATTATTTGGATGAGTAGAATTTTCCAAGTCTCTGGGTAGGCA 2775
Db 526 -----PheArgPheGly-----GlyGlyGly 532
QY 2774 TTTTGGGGGGCCAGA-----CCCCAGGAGAGAGAGATTTCTGGCAATGATC 2730
Db 533 AspAlaGlySerLysGlyProMetValSerAlaGlnGluSerGlnAlaGlnAlaIleLeu 552
QY 2729 AGCCCAATGACCACTATCTCAGGGGA-----CCT 2700
Db 553 GlnGlnAlaArgLeuAlaLeuArgGlyProAlaGlyProMetGlyLeuThrGlyArgPro 572
QY 2699 GATTGTTGGGATCCCCCACCCTACCCAAATATTAGACACACACAGAAAGGTAGCAA 2640
Db 573 GlyProMetGly-ProProGlySerGlyGlyLeuLysGlyGluProGlyAspMetGlyPr 592
QY 2639 TGGATTCCCTTCTACTTGTGTTAAATAAATAGTTAAATATTAAATGCTGTCTCTGT 2580
Db 592 oGlnGlyProArg-----GlyVa 598
QY 2579 GATGCAACAGAGGACCAACAGGCCACATCCTGATAAAGGTAAAGGGGTGGATCA 2520
Db 598 lGlnGlyProGlyProThrGlyLysProGlyArgArgGlyArgAlaGlySerAsp-- 617
QY 2519 GCAAAAGACAGTCTGTGGGCTGAGGGACCTGGTCTTGTGTGTGTGGCCCTCAGGACT 2460
Db 617 ----- 617
QY 2459 CTTCCCTTACAATAAGTCATATGTTCAATCCCATGGAGAGTGTTCATCCTTAGAAGC 2400
Db 617 ----- 617
QY 2399 TCCCATGCAAGAGCTACATTAAACGAGCTGAGGCTTAAGGGCTTAGAGATGGGAACC 2340
Db 618 -----GlyAlaArg---GlyMetPr 623
QY 2339 AGGTGACTGAGTTTATTTCAGCTCCGCAAAAACCCCTTCTAGTGTGTCTCAACTAGGAGG 2280
Db 623 oGlyGln-----ThrGlyProLysGly---AspArgGlyPheAspGlyLeuAlaGl 639
QY 2279 CTAGCTGTTAACCTTGAGCCCTGGGTATCCACCTGCGAGGTCGCCCGCATTCACGTGCATG 2220
Db 639 y-----LeuProGlyGluLysGlyHisArgGlyAspPr 650
QY 2219 GAGCCCTTCTGGCTCCCTGCTGATATAAGTCCAGACTGAAACCCCTTGGAAAGCCCTCCAGTC 2160
Db 650 oGlyProSerGlyProPro----- 656

```

Db 623 oGlyGln-----ThrGlyProLysGly---AspArgGlyPheAspGlyLeuAlaG1 639  
QY 2279 CTAGCTGTAAACCTCGAGCTGGGTAAATCCAGCTGCAGAGTCCCGCATTCACATGTCATG 2220  
Db 639 y-----LeuProGlyGluLysGlyHisArgGlyAspPr 650  
QY 2219 GAGCCCTTCTGCTCCCTGTATATAGTCCAGACTGAACCCCTTGGAAAGCCCTCCAGTC 2160  
Db 650 oGlyProSerGlyProPro----- 656  
QY 2159 AGGCAGCCCTAGAGACTGGGAGAGAGG-----AGAGGAC 2124  
Db 657 -GlyLeuProGlyAspAspGlyGluArgGlyAspGlyGluValGlyProArgGlyLe 676  
QY 2123 GCGCCAGCCCCAGCTGTGCAG-----CTACGCACCTTCAGCAGCAGAGGTGGCAGCAG 2070  
Db 676 uProGlyGluProGlyProArgGlyLeuLeuGlyProLysGlyProProGlyProG1 696  
QY 2069 AGAGCCACATTTACTTTGGCAGCAACAGNAAGTGGCGGCGCCAGCCCGGAGCCCATGGGGC 2010  
Db 696 yProPro-----Gly-ValThrGlyMetAspGlyGlnProGly---ProLysGlyA 712  
QY 2009 TAACAGGAGCGGGAGCTGGGACCCAGTGGAGCGCCCTCCACCCCAATGTGTGGAAG 1950  
Db 712 snValGly-----ProGln-GlyGluProGlyProPro----- 722  
QY 1949 TTTCTAGCTGAGTATTTGGCCAAAGTCGCTTTGTCAATPACTACCTGTGTAGCAAAAGT 1890  
Db 722 ----- 722  
QY 1889 AAATGGGACACAGCCAGGCTCGGCGAGACACCATATAGGAGTGCACAGACTGGCTGA 1830  
Db 723 GlyGlnGlnGlyAsnProGlyAlaGlnGlyLeuProGlyProGlnGlyAlaIleGlyPro 742  
QY 1829 GCTGGACAATGAGCCATAAACAGGATGGGCGCACCTGGGACAGCAGGAGGCACTAT 1770  
Db 743 ProGlyGluLysGlyPro-----LeuGlyLysProGlyLeuProGly-----Met 757  
QY 1769 CCAGATGGCGAGTTCAGGAGGATGCCCGGCGGAGAACCTCGCTGGCTGGTGGCT 1710  
Db 758 ProGlyAlaAspGlyPro-----Pro-GlyHisProGlyLysGluGly-- 771  
QY 1709 CACCCACACACACGTACGGAGACATCACAGGCAGAGCGCCGACAGCGCGGTGGAG 1650  
Db 772 -----ProProGlyGluLys-GlyGly 779  
QY 1649 GTGGAGAGGCCACTGCTCCAGCACCAGCTGCTCCATTAGG----- 1606  
Db 779 lngly-----ProProGlyProGlnGlyProIleGlyTyProGlyProArgG 795  
QY 1605 -----AAGGAGCTCCAGGCTTA-----GGCCTGGCAGGAAAGCTGTCA 1566  
Db 795 lyValLysGlyAlaAspGlyIleArgGlyLeuLysGlyThrLysGlyGluLys-GlyGlu 814  
QY 1565 TCAGGCTGTCTCTACTGCTAGCACCCTCAGTGTCCCTCGGTATTTGGCA-----GGA 1512  
Db 815 AspGlyPhePro-----GlyPheLysGlyAspMetGlyIleLysGlyAspArgGly 831  
QY 1511 ACACCTGCTTCCCGTGGTAGAGGAGG-----CCAGTGTGTAGGGCA 1467  
Db 832 GluileGlyProProGlyProArgGlyGluAspGlyProGluGlyProLysGlyArgGly 851  
QY 1466 GGATCTCAGGCGGTGAGAGGTGAACCCGTGAGGGCGGTGAGCTGTCTACCCAGGCCCA 1407  
Db 852 GlyProAsnGly-----AspProGlyPro 859  
QY 1406 CACTGTGGGAGGAGCATGTGGCAGCCAGCAGGAAAGCTGCCACACTGGCCAAAT 1347  
Db 860 LeuGlyProThrGly-----GluLysGlyLysLeu----- 869  
QY 1346 AGACTGCTGAGTCCGGAATGCTGCACAGCGGTCTCCATGACACAGAGAGACAGG 1287  
Db ----- 1287

Db 870 -----GlyValProGlyLeuPro-----GlyTyProGly 879  
QY 1286 AGATGGCGCACTCCAGGAACAGCC-----CCAGGCTGCCATCCGAAGCCCTTCATCAT 1233  
Db 880 ArgGlnGlnProLysGlySerIleGlyPheProGlyPheProGlyAlaAsnGlyGluLys 899  
QY 1232 AGTGTCTCCGGG---CCTCGGTGCGCGGTCTAGCTCTGGCAGCGCCCTGGTACAGCCCT 1176  
Db 900 GlyGlyArgGlyThrProGlyLysProGlyProArgGlyGlnArgGlyProThrGlyPro 919  
QY 1175 CGCCCAAGAAATCCGTGTAAACAGAGCTGAAGTCTATAGTGCATCCAGTCCACAGCT 1116  
Db 920 ArgGlyGluArgGly-----ProArgGlyIleThrGly 930  
QY 1115 CACCCAGGAAGACCGCGGCGGATGCGGCGGATGCGGCGAGCAGCAGCTGGTCCAGCCGG 1056  
Db 931 LysProGlyProLysGly-----AsnSerGlyGlyAspGly 942  
QY 1055 GAAGCAGGCGCCAGGTTCCCGAAAGCCAGCGGCGCGCATGGACAGCAGTGGGGG 996  
Db 943 ProAlaGlyProProGlyGluArgGlyProAsnGlyPro----- 955  
QY 995 ACAAGAGGCGCGCAGACGCCCTCTCTGCTGGCTGGGCGGCGGCGCTGCTCTCAG 936  
Db 956 -----GlnGlyProThrGlyPhePro-----GlyProLysGlyProGly 969  
QY 935 CCACCAGCAGTGTGGCTGTCTACGCGGTGAGGAAGATGAGGTGAGCAGCGCCAAAGAGC 876  
Db 970 ProProGly-----LysAspGlyLeuProGlyHisProGlyGlnArgGly 984  
QY 875 ACT-----CTCTGGGTGCCAGGT-----AGG 852  
Db 985 GluThrGlyPheGlnGlyLysThrGlyProProGlyProProGlyValValGlyProGln 1004  
QY 851 GGSCCA-----GGCAGCTGGTGTCCCACTCAATGG 822  
Db 1005 GlyProThrGlyGluThrGlyProMetGlyGluArgGlyHis----- 1018  
QY 821 CAGGCAGGAGGTAGCCAGCAGCGCCCAAGACTGATCATGAAGCATAGACAGAGTAGG 762  
Db 1019 -----ProGlyProProGlyProGlyGlu-----GlnGlyLeu 1030  
QY 761 CCGGCGCAGCTGGT-----CCGGTCCC-----GGAAGAGGTGAG 726  
Db 1031 ProGlyAlaIleGlyLysGlyThrLysGlyAspProGlyProAlaGlyLeuProGly 1050  
QY 725 AGACAGGCGCTCCAGTGGAGTGAAGCAGCACCTGGCCACAGAGTCCAGCAGCCCA 669  
Db 1051 LysAspGlyProProGlyLeuArgGlyPheProGlyAspArgGlyLeuProGlyProVal 1070  
QY 668 -----CGCCAGGATGAGCAGTGCAGCTCCAGG 639  
Db 1071 GlyAlaLeuGlyLeuLysGlySerGluGlyProProGly-----ProProGlyPro 1087  
QY 638 GCTGGGATCCGGGACACAGCCCTGTAGCCAGCGCGCTTGGGATGAGAAAGAGC 579  
Db 1088 AlaGlySerProGly-----GluArgGly 1095  
QY 578 TCAGCAGGATGCCCAAGAGCAGTGCACAGATGAAGGCGCGGCGGCGCATAGCTCCAC 519  
Db 1096 ProAlaGly----- 1098  
QY 518 GGCAGTGTCTACTGGCTGAGCCCTAGAGCGGGGACACAGACAGCCAGCCAGCTGACCAA 459  
Db 1099 AlaAlaGly-----ProIleGlyIleProGlyArgProGlyPro----- 1111  
QY 458 TGCCCGCAGCACCATTGTCGAATCTCTCTACCCCACTTCCAGCAGCAGCAGCGCA 399  
Db 1112 -----GlnGlyProProGlyProAlaGlyGluLysGly 1122  
QY 398 CATAGTGTATGCTGCGGCCAAACACACTCCAGGCCAAAGGTAGCAGGTGTACACCA 339  
Db 1123 -----ValProGlyGluGlyProGlnGlyProAla 1133

Db 1208 eu-----PheGlyGlnLysGlyAspLcylGlySerArgGly-----PheProG 1222

Qy 383 CGGCCAACACACTCCAGC-----CAAGGTTAGCAGGTGACACCAACAGAGCTGGCTT 327

Db 1222 ly-ProProGlyProValGlyLeuGlnGlyLeuProGlyProProGlyGlnLysGlyGlu 1241

Qy 326 TCCGGT-----GCCGACAGCAGCGGC 306

Db 1242 ThrGlyAspValGlyGlnMetGlyProProGlyProProGlyProArgGlyProSerGly 1261

Qy 305 TCACCCACACCTCTGACCATAGTGGCCA----- 275

Db 1262 AlaProGlyAlaAspGlyPro-GlnGlyProProGlyGlyIleGlyAsnProGlyAlaVa 1281

Qy 274 -GGCGGTAGGCTCAGGGGCGCTCAGGCACTCCAGAACTCTTCGTCCTGCGC----- 221

Db 1281 lGlyGlnLysGlyGluProGlyGlnAlaGlyAspProGlyLeuProGlyGlnGlyGlyPr 1301

Qy 220 -----TCTGCTCCAGAAAGCTCGCGCCTC 198

Db 1301 oLeuGlyProLysGlyLcylArgGlyGlnLysGlyGluAlaGlyProSerGlyAlaAlaGl 1321

Qy 197 TCTCTCTGCTGCGC-----CACTGCTAGGAATCAGCCAGCGCCCA 153

Db 1321 yProProGlyProLysGlyProProGlyAspAspGlyProLysGlySerProGly----- 1339

Qy 152 TTTCTGCGCAGCCTTTGGT-----GCCGTCAGCTTCTCAGCCCATGCTCAA 105

Db 1340 -----ProValGlyPheProGlyAspProGlyProProGlyGluPro-----G 1354

Qy 104 CACCTGCTGCTGTGGGCACTCAGTGGGACAGCTCTCATCACTCAGATCTCTGCGCGAG 45

Db 1354 y-ProAlaGlyGlnAspGlyProProGlyAspLysGlyAspGlyGluProGlyGlnT 1374

Qy 44 GCGCGCGCTGTACCCGGA 25

Db 1374 hrGly-----SerProGly 1378

RESULT 28

Q9J103

ID Q9J103 PRELIMINARY; PRT: 1840 AA.

AC Q9J103;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Alpha 1 type V collagen.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE DAWLEY;

RX MEDLINE=20428740; PubMed=10852920;

RA Chernousov M.A., Rotherbum K., Tyler W.A., Stahl R.C., Carey D.J.;

RT "Schwann cells synthesize type V collagen that contains a novel alpha

RT 4 chain. molecular cloning, biochemical characterization, and high

RT affinity heparin binding of alpha4(V) collagen.";

RL J. Biol. Chem. 275:28208-28215(2000).

DR EMBL: AF272662; AAF76433.1;

DR InterPro: IPR000087; Collagen.

DR InterPro: IPR000885; Fib\_collagen\_C.

DR InterPro: IPR001791; Laminin\_G.

DR InterPro: IPR001230; Prenyl\_site.

DR InterPro: IPR000508; SigPase.

DR InterPro: IPR003129; TSPN.

DR Pfam: PF01410; COLFI; 1.

DR Pfam: PF01391; Collagen; 17.

DR Pfam: PF02210; TSPN; 1.

DR ProDom: PD000007; Collagen; 2.

DR ProDom: PD02078; Fib\_collagen\_C; 1.

DR SMART: SM00038; COLFI; 1.

DR SMART: SM00282; LamG; 1.

DR SMART: SM00210; TSPN; 1.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN\_1.

DR PROSITE: PS00761; SPASE\_I\_3; UNKNOWN\_1.

KW Collagen.

SQ SEQUENCE 1840 AA: 183986 MW: AD38F5FF86B923C CRC64;

Alignment Scores:

Pred. No.:	1.99e-13	Length:	1840
Score:	335.00	Matches:	290
Percent Similarity:	28.9%	Conservative:	47
Best Local Similarity:	24.94%	Mismatches:	356
Query Match:	5.38%	Indels:	470
DB:	11	Gaps:	64

US-09-759-143-110 (1-3410) x Q9J103 (1-1840)

Qy 3068 GGTCTGTCAGTACCTCCAAACAGAGGTTGGAGCTGGTGGGAAAGTTGGGGTAGGGGA 3009

Db 433 GlyProGlyMetProAlaAsnGlnAspThrIleTyrGluGlyIleGlyGlyProArgGly 452

Qy 3008 AAGTTGGGG---GTAGGGGAAATTTGGGCAGTGCCTTCATCAGCCAGTCTCAGAGAGA 2952

Db 453 GlnLysGlyGlnLysGlyGlu-----ProAlaIleIleGluProGlyMet---Leu 468

Qy 2951 GTAGAGGGG----- 2940

Db 469 IleGluGlyProProGlyProGluGlyProAlaGlyLeuProGlyProGlyThrThr 488

Qy 2939 GGAAGTGGGGGAACACGCTGGCCAGAGAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880

Db 489 GlyProThrGlyGlnMetGlyAspProGlyGluArgGlyProProGlyArgProGlyLeu 508

Qy 2879 CCGTAAGGC-----CCACCTCTACTCTCTTCAACACCTAACCTTGGG 2835

Db 509 ProGlyAlaAspGlyLeuProGlyProGlyThrMetLeuMetLeuPro----- 525

Qy 2834 TAACAGCATTTGGAATATCATTTGGATGAGTAGAATTTCCAGGTCCTGGGTAGGCA 2775

Db 526 -----PheArgPheGly-----GlyGlyGly 532

Qy 2774 TTTTGGGGGCCAGG-----CCCGAGGAGAGAGATCTGGCAATGATC 2730

Db 533 AspAlaGlySerLysGlyProMetValSerAlaGlnGluSerGlnAlaGlnAlaLeu 552

Qy 2729 AGCCCAATGACCATCTCTCAGGGA-----CCT 2700

Db 553 GlnGlnAlaArgLeuAlaLeuArgGlyProAlaGlyProMetGlyLeuThrGlyArgPro 572

Qy 2699 GATTGTGGGGATCCCCACCCCTACCCCAATATTAGACACCAACACAGAAAAGCTAGCAA 2640

Db 573 GlyProMetGly-ProProGlySerGlyGlyLeuLysGlyGluProGlyAspMetGlyPr 592

Qy 2639 TGGATTCTCTCTTACTTTGTTAAATAAGTTAAATATTTAAATGCTGCTCTCTGT 2580

Db 592 oGlnGlyProArg-----GlyVa 598

Qy 2579 GATGGCAACAGAGGACCAACAGGCCATCTGTATAAAAGTTAGAGGGGGTGTGATCA 2520

Db 598 lGlnGlyProProGlyProThrGlyLysProGlyArgArgGlyArgAlaGlySerAsp-- 617

Qy 2519 GCNAAAGACAGTGTGTGGGTGAGGGGACCTGCTTCTGTGTGTGGCCCTCAGGACT 2460

Db 617 ----- 617

Qy 2459 CTTCCCTACAAATAAGTCATATGTTCAAAATCCCATGGAGAGTGTTTCATCCTAGAAAC 2400

Db 617 ----- 617

Qy 2399 TCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGGGAACC 2340

Db 618 -----GlyAlaArg---GlyMetPr 623

Qy 2339 AGTGACTGAGTTATTTCAGCTCCCAAAACCCCTCTCTAGGTGTCTCTCACTAGGAGG 2280



Db 900 almetValGlyProGlyAlaLysGlyGluLysGlyAlaProGlyAspLeuAlaGlyA 920  
 QY 308 -----GGCTCACCACACGCTCTGGACCATAGTGG----- 278  
 Db 920 spLeuLeuGlyGluProGlyAlaLysGlyAspArgGlyLeuProGlyProArgGlyGlu 939  
 QY 277 -----CCAGGCGGGTAGGCTCAGGCGCGCGTTCAGGC 245  
 Db 940 LysGlyGluAlaGlyArgAlaGlyGluProGlyAspProGlyGluAspGlyGlnLysGly 959  
 QY 244 ACTCAGAACTGCTGCTCGGC-----TCTGCTCCA 212  
 Db 960 AlaProGly-----ArgLysGlyLeuLysGlyGluProGlyIleGlyValGlnGlyPro 977  
 QY 211 GAAGCTGGGCTCTCTCTCTGCTGCGC----- 182  
 Db 978 ProGlyProThrGlyProGlyProGlyMetLysGlyAspValGlySerProGlyAlaProGly 997  
 QY 181 -----CAACTGCTAGG-----AATCAGCCAGGCGCCCAT 152  
 Db 998 ValValGlyPheProGlyGlnThrGlyProArgGlyGluThrGlyGlnProGly-ProVa 1017  
 QY 151 TTCTGCCAGCCTTTGGTG-----CCGTCCAGCT 122  
 Db 1017 lGlyGluArgGlyLeuAlaGlyProProGlyArgGluGlyAlaProGlyPro----- 1034  
 QY 121 TCTCAGCCATGCTCAACACCTGCTGCTGTGGG-----CACCTCAGTGGGACAGTC 68  
 Db 1035 -LeuGlyProProGlyProGlySerValGlyAlaProGlyAlaSerGlyLeuLysGly 1054  
 QY 67 TCATCAGTCAATGCTGCG-----CGAGCGCGCGGTGTCCACCGG 26  
 Db 1054 yAspLysGlyAspProGlyThrGlyLeuProGlyProArgGlyGluArgGlyGluProGly 1074  
 QY 25 A 25  
 Db 1074 y 1074

RESULT 27  
 O88207 PRELIMINARY; PRT; 1838 AA.  
 AC O88207  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 GN COL5A1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98250615; PubMed=9582436;  
 RA Wu Y.L., Sumiyoshi H., Khaleduzzaman M., Ninomiya Y., Yoshioka H.;  
 RT "cDNA sequence and expression of the mouse alpha1(V) collagen gene.";  
 RL Biochim. Biophys. Acta 1397:275-284(1998).  
 DR EMBL: AB009993; BAA28786.1;  
 DR MGD: MGI:88457; Col5a1.  
 DR InterPro: IPR000087; Collagen.  
 DR InterPro: IPR000885; Fib\_collagen\_C.  
 DR Pfam: PF01391; Collagen; 18.  
 DR InterPro: IPR001791; Laminin\_G.  
 DR InterPro: IPR001230; Prenyl\_site.  
 DR Pfam: PF01410; COLFI; 1.  
 DR Pfam: PF01391; Collagen; 18.  
 DR ProDom: PD000007; Collagen; 2.  
 DR ProDom: PD002078; Fib\_collagen\_C; 1.  
 DR SMART: SM00038; COLFI; 1.  
 DR SMART: SM00282; LamG; 1.  
 DR SMART: SM00210; TSPN; 1.  
 DR PROSITE: P500294; PRENYLATION; UNKNOWN\_1.

KW Collagen. 1838 AA; 183691 MW; 7A520B23D1851783 CRC64;  
 SQ SEQUENCE  
 Alignment Scores:  
 Pred. No.: 1,85e-13 Length: 1838  
 Score: 335.50 Matches: 301  
 Percent Similarity: 29.74% Conservative: 60  
 Best Local Similarity: 24.79% Mismatches: 395  
 Query Match: 5.39% Indels: 458  
 DB: 11 Caps: 68

US-09-759-143-110 (1-3410) x 088207 (1-1838)  
 QY 3068 GGTCTCTGACAGTCTCAACACAGGTTGTGGAGCTGTGGGAAAGTGTGGGTAGGGA 3009  
 Db 431 GlyProGlyMetProAlaAsnGlnAspThrIlePheGluGlyIleGlyGlyProArgGly 450  
 QY 3008 AAGTTGGG---GTAGGGAAATTTGGGCGAGTCTTCATCAGCCAGCTCCTAGAGAGA 2952  
 Db 451 GluLysGlyGlnLysGlyGlu-----ProAlaIleIleGluProGlyMet---Leu 466  
 QY 2951 GTAGAGGG-----AGT 2940  
 Db 467 IleGluGlyProGlyProGlyProGluGlyProAlaGlyLeuProGlyProGlyThrThr 486  
 QY 2939 GGAAGTGGGGGGAACACAGGCTGGGCCAAGAGAGAGGGTGTAGGGAGCGCTTGAGA 2880  
 Db 487 GlyProThrGlyGlnMetGlyAspProGlyGluArgGlyProGlyArgProGlyLeu 506  
 QY 2879 CCTGAGCC-----CCACCTCTACCTTCTTCAACACCCCTAACCTTGGG 2835  
 Db 507 ProGlyAlaAspGlyLeuProGlyProGlyThrMetLeuMetLeuPro----- 523  
 QY 2834 TAACAGCATTTGGAATTTATCATTTGGGATGAGTAGAATTTCCAAAGGTCTCTGGTTAGGCA 2775  
 Db 524 -----PheArgPheGly-----GlyGlyGly 530  
 QY 2774 TTTTGGGGGCCAGA-----CCCCAGGAGAGAGATTTTGGCAATGATC 2730  
 Db 531 AspAlaGlySerLysGlyProMetValSerAlaGlnGlnSerGlnAlaGlnAlaLeu 550  
 QY 2729 AGCCCAATGACAGCTATCTCAGGGGA-----CCT 2700  
 Db 551 GlnGlnAlaArgLeuAlaLeuArgGlyProAlaGlyProMetGlyLeuThrGlyArgPro 570  
 QY 2699 GATTGTGGGATCCCCACCCCTACCCAAATATTAGACACCAACACACAGAAAGTAGCAA 2640  
 Db 571 GlyProMetGly-ProProGlySerGlyGlyLeuLysGlyGluProGlyAspMetGlyPr 590  
 QY 2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTAAATGCCCTGTCTCTGT 2580  
 Db 590 oGlnGlyProArg-----GlyVa 596  
 QY 2579 GATGGCAACAGAGAGGACCAACAGCCACATCTCTGATATAAGGTTAAGAGGGGTGGATCA 2520  
 Db 596 lGlnGlyProGlyProThrGlyLysProGlyArgArgGlyArgAlaGlySerAsp-- 615  
 QY 2519 GCAAAAAGACAGTCTGTGGGCTGAGGGGACCTGTTCTTGTGTGTTGCCCTCAGGACT 2460  
 Db 615 ----- 615  
 QY 2459 CTTCCTCCCTACAAATAAGTTCATATGTTCAATCCCATGAGAGAGTGTTCATCCTAGAAC 2400  
 Db 615 ----- 615  
 QY 2399 TCCCATGCAAGAGCTACATTAAAGAGAGCTGCAGGTTAAGGGCTTAGAGATGGGAACC 2340  
 Db 616 -----GlyAlaArg---GlyMetPr 621  
 QY 2339 AGGTGACTGAGTTTATTCAGTCCCAAAACCCCTTCTCTAGGTGTGTCTCACTAGGAGG 2280  
 Db 621 oGlyGln-----ThrGlyProLysGly---AspArgGlyPheAspGlyLeuAlaGl 637





```

Db 364 valLysArgLeuTrpGlyIleValAsnPheLeuLeu-----AlaIleCysLeuGly 380
QY 1400 CACAGTGGCGGTGGT-----1417
Db 381 LeuThrValLeuValThrLysLeuAlaGlnHisSerArgGlnTyrAlaProGlyThrGly 400
QY 1418 -----ACAGCTTCAGCCGCCCTCACCGGTTTC 1444
Db 401 AlaLeuGlyAspProLeuProSerGluGlyIleLysAlaGlyAlaLeuThrLeuPhe 420
QY 1445 ACCTTCTAGCCCTGCAGATC-----CTGCCCTACACACTGCCCTCCCTC 1489
Db 421 SerValLeuGlyValProLeuAlaIleThrTyrSerIleProPheAlaLeuAlaSerIle 440
QY 1490 TACCACCGGAGAGCAGGTGCTCTGCCCAATACCAGGGGACACTGGAGTCTACG 1549
Db 441 Phe-----SerSerThrSerGlyAlaGly 448
QY 1550 ACTGAGCAGACGCTG-----ATGACCAGC 1573
Db 449 GlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValIleProGlnMetPheValSer 468
QY 1574 TTCCTGCCAGCCCTAAGCCTGGA-----GTCCTCTTCCCTAAT 1612
Db 469 ValLeuSerGlyProTrpAspAlaLeuPheGlyGlyGlyAsnLeuProAlaPheValVal 488
QY 1613 GGACAGCTGGTGGTGGAGGAGTGGC-----CTGCTCCACCTCCACCC 1657
Db 489 GlyAlaValAlaAlaLeuAlaSerGlyIleLeuSerIleIleLeuLeuProSerProPro 508
RESULT 26
Q60444
ID Q60444 PRELIMINARY; PRT: 1549 AA.
AC Q60444;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Type VII collagen (Fragment).
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE=93271985; PubMed=8499916;
RA Greenspan D.S.;
RT "The carboxyl-terminal half of type VII collagen, including the non-
RT collagenous NC-2 domain and intron/exon organization of the
RT corresponding region of the COL7A1 gene."
RL Hum. Mol. Genet. 2:273-278(1993).
DR EMBL; L06863; AAA36968.1; -
DR HSSP; P00981; 1DTH.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF01391; Collagen; 22.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR PRINTS; PR00759; BASICTPASE.
DR ProDom; PD000007; Collagen; 9.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS02079; BPTI_KUNITZ_2; 1.
KW Collagen; Serine protease inhibitor.
FT NON_TER 1
SQ SEQUENCE 1549 AA; 148117 MW; 5D646391E0C29292 CRC64;

```

```

Alignment Scores:
Pred. No.: 1.8e-13 Length: 1549
Score: 335.50 Matches: 324
Percent Similarity: 30.87% Conservative: 61
Best Local Similarity: 25.98% Mismatches: 419

```

```

Query Match: 5.39% Indels: 444
DB: 11 Gaps: 73
US-09-759-143-110 (1-3410) x Q60444 (1-1549)
QY 3039 GGAGCTGTGGGAAAGTTGGGTAGGGAAAGTTGGGTAGGGGAAATTTTGGGCAG 2980
Db 36 GlySerProGlyProGlnGlyProAlaGlyArgAlaGlyGluLysGlyGluLysGlyAsp 55
QY 2979 TGCCTTCAT-----CAGCCC-----2965
Db 56 CysGluAspGlyAlaProGlyLeuProGlyGlnProGlyAlaProGlyGluProGlyLeu 75
QY 2964 -----AGTCCTAGACAGAGTAGTACAGGGGAGTGAAGTGGGGGAA 2926
Db 76 ArgGlyThrProGlyIleThrGlyProLysGlyAspArgGlyGln-----ThrGlyThr 93
QY 2925 CCAGCTGGGCGCAAGAGAGAGGGTGTAGGAGCCGTTGAGACTGAAGCCGCCAC 2866
Db 94 ProGly-GluProGlyGluLysGlyGluArgGlyProPro-GlyProValGlyProGln 113
QY 2865 CTCTACCTTCTTCAACACCCTAACCTTGGGTAAACAGCATTTGGAATATCATTTGGAT 2806
Db 113 LyLeuProGly-----ValAlaGly-----HisProGly- 122
QY 2805 GAGTAGAATTTTCAAGGTCTGTGGTGTAGGCATTTTGGGGGCGCAGACCCAGGAGAA 2746
Db 123 --ValGluGlyProGlu-----GlyProGlyProAlaGlyArgA 136
QY 2745 GATTCTGGCAATGATCAGCCCAATGACAGCATATCTCAGGGGACCTGATTTGGGGATC 2686
Db 136 rgGlyGluLysGlyGluProGlyArgPro-----GlyAsp 148
QY 2685 CCCACCCCTACCCA---AATATTAGACACCAACAGAAAAGCTAGCAATGGATTCCTT 2629
Db 148 roAlaValGlyProGlyGlyAlaGlyAlaLysGlyGluLysGlyAspAlaGlyLeuPro- 167
QY 2628 CTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCTGTCTGTGTGATGGCAACAG 2569
Db 167 -----167
QY 2568 AAGCACCAACAGGCCACATCTGTATAAAGTTAGAGGGGGTGGATCAGCAAAAGACA 2509
Db 168 --GlyProArgGlyAlaAlaGlyIleLysGlyGluGlnGly-----180
QY 2508 GTGCTGTGGGTGAGGGGACCTGCTTCTGTGTGCTGCCCTCAGGACTCTTCCCTTACA 2449
Db 181 -----ProGlyLeu-----184
QY 2448 AATAAGTCATATGTTCAATCCCATGGAGAGTCTTTTCATCTAGAAACTCCCATGCAAG 2389
Db 185 -----AlaLeuProGlyAspProGlyProLys-----193
QY 2388 AGCTACATTAAACCAAGCTGCAGTTAAGGGCTTAGAGATGGGAAACAGGTGACTGAG 2329
Db 194 -----GlyAspProGlyAsp-----198
QY 2328 TTTATTACGCTCCCAAAAACCTTCTCTAGTGTGTCTCAACTAGGAGGTAGCTGTATA 2269
Db 199 -----ArgGlyPro---IleGlyLeuThrGlyArgAlaGly-----P 210
QY 2268 CCCTCAGCCTGGGTAAATCCACCTGCAGATGCCCGCATTCCTCAGTGCATGGAGCCCTTCTG 2209
Db 210 roThrGlyAspSerGlyProProGlyGluLysGlyAspProGlyArgProGlyProProG 230
QY 2208 GCCTCCCTGTATAGTCCAG-----ACTGAACCCCTTGAAGGCC 2167
Db 230 ly---ProValGlySerArgGlyArgAspGlyGluValGlyGlyValGlyAla 249
QY 2166 TC-----CAGTCAGGAGCCCTAGAGACTGGGGAGAGAGG-----AGAGGGAGCC- 2120
Db 249 snProGlyAspProGlyLeuProGlyLysAlaGlyGluArgGlyLeuArgGlyAlaProG 269

```







257 y-----LeuProGlyMetLysHisArgGlyPheSerGlyLeuAspG1 272  
Db  
Qy 1892 AGTAATGGGACAGACAGCCAGCCCTCCGGCAGACACACCATATAGGAGCTGACAGATGGC 1833  
Db  
272 yAlaLysGlyAspAlaGlyProAlaGlyProLysGlyGluProGlySerProGly----- 290  
Qy 1832 TGAGCTGGACAATGGAGCCCAATAACAGGATGGGCCA-----CCTGGCAGCA 1782  
Db  
291 -----GluAsnGlyAlaProGlyGln--MetGlyProArgGlyLeuProGlyGluArg 307  
Db  
1781 GGAAGGCACTATCCAGGATGGAGGTCAGAGTCCAGGAGATGCCCGGCCCGGAA----- 1731  
Qy  
308 GlyArg-----ProGlyAlaProGlyProAlaGlyAlaArgGlyAsnAspGlyAlaThr 325  
Db  
1730 -----CCACCTGGCTCGGTGGCTCACCACCCACCA----- 1696  
Qy  
326 GlyAlaAlaGlyProProGlyProThrGlyProAlaGlyProProGlyPheProGlyAla 345  
Db  
1695 -----CCTACGGAGACATCACAG-----GCAGAGGCCCGCAGAGC 1660  
Qy  
346 ValGlyAlaLysGlyGluAlaGlyProGlnGlyProArgGlySerGluGlyProGlnGly 365  
Db  
1659 GCGGGTGGAGTGGGACAGGCCACTCCTCCAGCAGCCACCGTGTCCATTAGG-----AAG 1603  
Qy  
366 ValArgGly-----GluProGlyProProGlyProAlaGlyAlaAlaGlyProAla 382  
Db  
1602 GGAGCTCCAGCTTAGG-----CCTGGCAGGAGCTGCTCATCAGGCTGTCTCACTG 1549  
Qy  
383 GlyAsnProGlyAlaAspGlyGlnProGlyAlaLys--GlyAlaAsnGlyAlaPro----- 400  
Db  
1548 CTAGCAGCTCCAGTGTCCCTCGGTATTGGGAGGAGAACACCTCTCTCCCGGTGTAG 1489  
Qy  
401 -----GlyIle-----AlaGlyAlaProGlyPheProGlyAlaAr 412  
Db  
1488 AGGAGGCCAGTGTAGGCGAGATCTGCAGGCTGAGAGGTGAACCCGTTAGGGCG 1429  
Qy  
412 gGly-----ProSerGlyPro--GlnGlyProGlyGlyProGlyProLysG 428  
Db  
1428 GCTGAAGCTGTACCACGGCCACACTCTGGCAGGAGCTGGCAGCCGCGCAGCAGG 1369  
Qy  
428 LysnSerGlyGluProGly---AlaProGlySerLysGlyAspThrGly--AlaLysGly 446  
Db  
1368 AAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCAGTGCAGTGCACAGCCGGTCC 1309  
Qy  
447 GluProGlyPro-----Val--GlyValGlnGlyProProGlyPr 459  
Db  
1308 ATGACCAG-----AGAAGACCAGGAGATGCCGACCTCCAGGAACAGCC---CCAGG 1258  
Qy  
459 oAlaGlyGluGlyLysArgGlyAlaArgGlyGluProGlyProThrGlyLeuProG1 479  
Db  
1257 CTGCCCATCCGAAGC-----CTTCATCATAGTGTCTC 1225  
Qy  
479 yProProGlyGluArgGlyGlyProGlySerArgGlyPheProGlyAlaAspGlyValAl 499  
Db  
1224 CGGGCTCGGTGGCGGTGAGCTCTGGGACGCGCTGTGTACAGCCCTCGCCACCAAA 1165  
Qy  
499 aGlyProLysGlyProAlaGlyGluArgGlySerProGlyProAlaGlyProLysGlySe 519  
Db  
1164 TCCGTGTAAACAGCGGTGAAGTTCATGAGTCCCATCCAGTGCACAGCTCAGCCACGAAG 1105  
Qy  
519 fPro----- 520  
Db  
1104 AGCCGGCGCAGGGTGGGGGATCGGCGATCGGACAGAGCTGGTACGCGGGGAAGCAGGGCG 1045  
Qy  
521 -----GlyGluAlaGlyArgProGlyGluAlaGlyLe 531  
Db  
1044 CCCAGT-----TCCGAAGCCAGCGGGCCCGGATGACAGCAGTGGGGC 997  
Qy  
531 uProGlyAlaLysGlyLeuThrGlySerProGlySerProGlyProAspGly----- 548  
Db  
996 GACAAGAGGGGGCGACAGCCCTCTCTGGCTCGGTGGGGCCCGCAGC----- 948  
Qy  
549 -----LysThrGlyProPro-----GlyProAlaGlyGlnAspG1 560  
Db  
947 -----CTGCTCTCTCAGCCACCA----- 930  
Qy  
560 yArgProGlyProProGlyProProGlyAlaArgGlyGlnAlaGlyValMetGlyPhePr 580  
Db  
929 -----GCAGTGTGGCTGCTACGAGGTGAGGAGATGAGGTGAGCAGGCCA 883  
Qy  
580 oGlyProLysGlyAlaAla-----GlyGluProGlyLysAlaGlyL 594  
Db  
882 ANAGGCACT-----CCTCTGGTGGTGGCCAGGT----- 855  
Qy  
594 uArgGlyValProGlyProProGlyAlaValGlyProAlaGlyLysAspGlyGluAlaG1 614  
Db  
854 -----AGGGGCCAGGCGACTGGTCTCCAGTCAATGCGCAGC 817  
Qy  
614 yAlaGlnGlyProProGlyProAlaGlyGluArgGlyGluGlnGlyProAl 634  
Db  
816 AGGA-----GGTAGCCAGGCGAGCCGCCAGAGACTGATCATG 781  
Qy  
634 aGlySerProGlyPheGlnGlyLeuProGlyProAlaGlyProProGlyGluAlaGlyL 654  
Db  
780 AAGGCATACAGACTAGGCTGGCGACAGTGT---CGGGTCCC-----GGAAGAGG 730  
Qy  
654 sProGlyGluGlnGlyValProGlyAspLeuGlyAlaProGlyProSerGlyAlaArgG1 674  
Db  
729 TCAGAGCAGGAGCGCTCCAGTGGAGTGAAGCACACCTGCCACAGAAAGTCCAGCAGCC 670  
Qy  
674 yGluArg---GlyPheProGlyGluArgGlyValGlnGlyProProGlyProAlaGlyPr 693  
Db  
669 AGCCCGAGTACAGCAGTCCCA-----GCTCCAGG 640  
Qy  
693 oArgGlyAlaAsnGlyAlaProGlyAsnAspGlyAlaLysGlyAspAlaGlyAlaProG1 713  
Db  
639 GGCTGGATCCGGGCAGCAGCAGCCCTGCTAGCAGCCGCTTGGGATGAGAAGAGG 580  
Qy  
713 yAla-----ProGlySerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArgG1 731  
Db  
579 CTCAGCAGGATGCCAAGGACAGTGCACAGTGAAGGCGCGCGGCGGCGCATAGCTCCA 520  
Qy  
731 yAlaAlaGlyLeuProGly-----ProLysGlyAspArgGlyAspAlaGlyProLysG1 749  
Db  
519 CGCCAGTGTCTACTGCTGAGCTAGCAGG---GGACACAGA----- 480  
Qy  
749 yAlaAspGly-----SerProGlyLysAspGlyValArgGlyLeuThrGlyProL1 766  
Db  
479 -----CCAGGCCAGCAGCTGGACCAATGCCCA----- 453  
Qy  
766 eGlyProProGlyProAla-----GlyAlaProGlyAspLysGlyGlySerGlyProse 784  
Db  
452 -----GCACCATGTCATGACCTTCTCTCTTACC 424  
Qy  
784 rGlyProAlaGlyProThrGlyAlaArgGlyAlaProGlyAspArgGlyGluProGlyPr 804  
Db  
423 CCACCTTCCAGCAGCAGGCGGCACATAGGTGATGCTCGGCCCAACACACACTCCAGG 364  
Qy  
804 oProGlyProAlaGlyPheAlaGly-----ProProG1 815  
Db  
363 C-----CAAAGGTAGCAGGTTGACCAGCAAGAGTGGCTTCCGGTCCCGC 316  
Qy  
815 yAlaAspGlyGlnProGlyAlaLysGlyGluProGlyAspAlaGlyAlaLysGlyAspAl 835  
Db  
315 AGCAGGCGGCTCACCCACAGCCCTCTGGACCATAGTGGCCA-----GGCGGT 268  
Qy  
835 a-----GlyProProGlyProAlaGlyPro--AlaGlyProProGlyProL1eGlyAsnV 853  
Db  
267 AGGCTCAGGGCGCTTCCAGGCATCCAGACTGCTCTGCTCTGCGCTCTGCT---C 214  
Qy  
853 alGlyAlaProGlyAlaLysGlyAla-----ArgGlySerAlaGlyProp 868  
Db  
213 CAGAAGCTCGGCTCTCTCTCTGCTGCTGCGCCCACTGCTAGATCAGCCAGCGGCC 154  
Qy  
868 roGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSerGlyA 888  
Db

3571 nProGly 3573

Db

RESULT 22

076045

ID O76045 PRELIMINARY; PRT; 1461 AA.

OT 076045;

DC 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Pro alpha 1(I) collagen.

DE COL1A1.

OS Homo sapiens (Human).

GN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

XP [1]

RP SEQUENCE FROM N.A.

RP MEDLINE=85130970; PubMed=2857713;

RA Chu M.L., de Wet W., Bernard M., Ramirez F.;

RT "Fine structural analysis of the human pro-alpha 1 (I) collagen gene.

RT promoter structure, AluI repeats, and polymorphic transcripts.";

RL J. Biol. Chem. 260:2315-2320(1985).

RL [2]

RP SEQUENCE FROM N.A.

RP MEDLINE=86329734; PubMed=2843432;

RA D'Alessio M., Bernard M., Pretorius P.J., de Wet W., Ramirez F.;

RT "Complete nucleotide sequence of the region encompassing the first

RT twenty-five exons of the human pro alpha 1(I) collagen gene.";

RL Gene 67:105-115(1988).

RL [3]

RP SEQUENCE FROM N.A.

RP MEDLINE=89025644; PubMed=3178743;

RA Tromp G., Kuivaniemi H., Stacey A., Shikata H., Baldwin C.T.,

RA Jaenisch R., Prockop D.J.;

RT "Structure of a full-length cDNA clone for the prepro alpha 1(I) chain

RT of human type I procollagen.";

RL Biochem. J. 253:919-922(1988).

RL [4]

RP SEQUENCE FROM N.A.

RP MEDLINE=91138770; PubMed=1995349;

RA Maatta A., Bornstein P., Penttinen R.P.;

RT "Highly conserved sequences in the 3'-untranslated region of the

RT COL1A1 gene bind cell-specific nuclear proteins.";

RL FEBS Lett. 279:9-13(1991).

RL [5]

RP SEQUENCE FROM N.A.

RP MEDLINE=92157916; PubMed=1787829;

RA Westerhausen A., Constantinou C.D., Pack M., Peng M.Z., Hanning C.,

RA Olsen A.S., Prockop D.J.;

RT "Completion of the last half of the structure of the human gene for

RT the pro alpha 1 (I) chain of type I procollagen (COL1A1).";

RL Matrix 11:375-379(1991).

RL [6]

RP SEQUENCE FROM N.A.

RP MEDLINE=96107942; PubMed=9443882;

RA Korkko J., Ala-Kokko L., De Paape A., Nuytink L., Earley J.,

RA Prockop D.J.;

RT "Analysis of the COL1A1 and COL1A2 genes by PCR amplification and

RT scanning by conformation-sensitive gel electrophoresis identifies only

RT COL1A1 mutations in 15 patients with osteogenesis imperfecta type I:

RT identification of common sequences of null-allele mutations.";

RL Am. J. Hum. Genet. 62:98-110(1998).

RL [7]

RP SEQUENCE FROM N.A.

RP MEDLINE=96107942; PubMed=9443882;

RA Korkko J., Earley J.J., Nuytink L., DePaape A., Prockop D.J.,

RA Ala-Kokko L.;

RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RA EMBL; AF017178; AA894054.2; -

DR InterPro; IPR000087; Collagen.

DR InterPro; IPR000885; Fib.Collagen\_C.

DR InterPro; IPR001007; VWF\_C.

DR Pfam; PF01410; COLFI; 1.

DR Pfam; PF01391; Collagen; 18.









Db	489	SerGlyValLeuAlaIleValLeuLeuProLysProSerLysAspAlaSerLysLeu	508
QY	1682	GATGCTCTCC 1690	1691
Db	509	SerLeuSer 511	512
RESULT 20			
Q9ESQ2			
ID	Q9ESQ2	PRELIMINARY;	PRT; 1691 AA.
AC	Q9ESQ2		
DT	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Type IV collagen alpha 5 chain.		
GN	COL4A5.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20536494; PubMed=10965041;		
RA	Saito K., Naito I., Seki T., Ohashi T., Kimura E., Momota R.,		
RA	Kishiro Y., Sado Y., Yoshioka H., Ninomiya Y.;		
RT	"Differential Expression of Mouse a5(IV) and a6(IV) Collagen Genes in		
RT	Epithelial Basement Membranes.";		
RL	J. Biochem. 128:427-434 (2000).		
DR	EMBL; AB041350; BAB13673.1; -.		
DR	MGD; MGI:88456; Col4a5.		
DR	InterPro; IPR000087; Collagen.		
DR	InterPro; IPR001442; ProcollagenC4.		
DR	Pfam; PF01413; C4; 2.		
DR	Pfam; PF01391; Collagen; 24.		
DR	ProDom; PD000007; Collagen; 3.		
DR	ProDom; PD003923; ProcollagenC4; 1.		
DR	SMART; SM00111; C4; 2.		
DR	Collagen.		
SK	SEQUENCE 1691 AA; 161823 MW; 81340DF192208FA CRC64;		
QW			
Alignment Scores:			
Pred. No.:	4,14e-14	Length:	1691
Score:	345.50	Matches:	274
Percent Similarity:	28.91%	Conservative:	55
Best Local Similarity:	24.08%	Mismatches:	369
Query Match:	5.38%	Indels:	440
DB:	11	Gaps:	57
US-09-759-143-110 (1-3410) x Q9ESQ2 (1-1691)			
Qy	129	CGCGACCAAGGGCTGGCAGAAATGGCGGCTGGCTGATTCCTAGGCAGTGTGGCGCAG	1888
Db	532	ProGlyProProGly-----AlaProGly---PheProGlySerLysGlyAsp	546
Qy	189	CAAGGAGGAGAGCGCCAGCTTCTGGACAGACGGCAGACGACGATCTCTGGAGTGCCT	248
Db	547	ProGlyAspValLeuLeuProGlyMetLysGlyAspLysGlyGluLeuGlyPhePro	556
Qy	249	GAACGGCCCTTACGCCCTACCGCTGGCCCACTATGTT-----	2888
Db	567	GlyAlaProGlyLeuProGlyLeuProGly-SerProGlyLysAspGlyLeuProGlyLe	586
Qy	289	-CCAGAGGCTGTGGTGAGCCG-----CCTGCTGGGCACCGGA	3256
Db	586	uProGlyProLysGlyGluProGlyGlyIleThrPheLysGlyGluArgGlyProProG	606
Qy	327	AAGCCAGCTTGTGGTGCACACCTTGAACCTTTGGCCTGGAGGTGTGTTGGCCGAC	386
Db	606	ySerProGlyLeuProGlyLeuProGlyLysnMetGlyProThrGly-----	621
Qy	387	GCATCAGCTATGTGGCGCCTCTGCTGCTGGAGAGTGGGGGTAGAGAGAAGTTCATGACCA	448
Db	622	-----ProValGlyPheGlyProProGlyProLysGlyGlu	634



QY 1897 AGCAAGTAATGGGACACAGACC-----AGGCCTGC 1865  
Db 150 -----Gly-ThrArgProProAlaProSerSerArgThrGlyGlnArg-ThrA 165  
QY 1864 GGCACACACATATAGGAGTGCACAGACTGCTGAGCTGGACATGGAGCCCATAAACAG 1805  
Db 165 rglInArgPro-----GlyHisProProGluArg 176  
QY 1804 GGATGGGCGCACCTGGGACA-----GCAGGAA 1778  
Db 176 lySerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgThrA 196  
QY 1777 GGCATATCCAGATGGCGAGTCCAGCAGATGCCCGCGCCG-----GAACAC 1727  
Db 196 rglInArgProGlyHisProProGluArgGlySerGlyProArgGlyThrArgProp 216  
QY 1726 CCGTGCCT-----CGTGGGCTCACCCACACCA 1698  
Db 216 roAlaProSerSerArgThrGlyGlnArgThrArgGlnArgProGlyHisProProG 236  
QY 1697 -----C 1697  
Db 236 luArgGlySerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnA 256  
QY 1696 AGCTAGCAGACATCACAGGAGGCGCCCGCAGAGCGCGGTGGAGTGGAGCAGGCC 1637  
Db 256 rglInArgGlnArg-ProGlyHisProProGluArg-GlySerGlyProArgGlyTh 275  
QY 1636 ACTGCTCCAGCAGCCAGCTGCCATTAGGAAG-----GGAGCTCC 1595  
Db 275 rArgProAlaProSerSerArgThrGlyGlnArgThrArgGlnArgProGlyHisPr 295  
QY 1594 A-----GGCTAGGCGCT---GGCAGGAAGCTGTGCATCAGGCTGTCTCCTACT 1550  
Db 295 oProProGluArgGlySerGlyProArgGlyThrArg-----307  
QY 1549 GCTAGCAGCTCCAGTGTCCCTCGTATTGGGAGGAAACACCTGCTTCTCCCGTGGTA 1490  
Db 308 -----ProProAlaProSerSerArgThrGlyGlnArgThr-ArgGlnArgProGlyH 325  
QY 1489 GAGGAGGCGAGTGTAGGCGAGGATCTGCAGGCTGAGAGGTGAACCGGTGAGGGC 1430  
Db 325 isProProGluArgGlySerGlyProArgGlyThrArg-----338  
QY 1429 GGCTGAAGTGTACACGCGGCACACTGTGGACAGCATGTGCACCGCAG-----1377  
Db 339 -----ProProAlaProSerSerArgThrGlyGlnArgThrArgGlnArgProG 355  
QY 1376 -----CCACAGGGAAGCTGCCACACTGGCCA---A 1349  
Db 355 lyHisProProGluArgGlySerGlyProArgGlyThrArgProProAlaProSerS 375  
QY 1348 ATAGACTGCTCAGTCCGGAATCGCTGCACGAGCGGTCCATACACAGAGAGAGACCCAG 1289  
Db 375 erArgThrGlyGlnArgThrArgGlnArgProGlyHisProProGluArgGlySerG 395  
QY 1288 GGAGATGGCGCACTGCAGGAACAGCCCGAGCTGCCATCCAGCGCTTCATCATAGTG 1229  
Db 395 ly-----ProArgGlyThrArgProProAlaProSerArg-----407  
QY 1228 TCTCCGGGCTCGGTGCGGCTCAGCTCTGGGACGCGCTGGTACAGCCCTCGCCAC 1169  
Db 408 ---ThrGlyGlnArgThrArg-----GlnArgProGlyHisProProGluA 423  
QY 1168 GAAATCCGCTGTAACACAGCGTGAAGTGCATGAGTGCATGAGTGCAGCTGCAGCTCAGCCAC 1109  
Db 423 rglSerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgT 443  
QY 1108 GAAGACCGCGCAGGCTGCGGGCATCGGCAGACACAGCTGGTGCAGCGGGAAGCAG 1049  
Db 443 hrArgGlnArgProGlyHisProProGluArgGlySerGly--ProArgGlyThrA 462  
QY 1048 GCGGCCAGGTTCCGGAAAGCCAAAGCGCGCGCATGACAGCAGTGGGCGGACAGGA 989

Db 462 gProPro-----AlaProSerSerArgThrGlyGlnArgThrArgGlnArgPr 478  
QY 988 GGGGGCGGACAGCCCTCTCTGCTGGCTGGTGGGGCCCGGCTCTCTCAGCCACAG 929  
Db 478 oGlyHisProProGluArgGlySerGlyProArgGlyThrArgProProAlaProSe 498  
QY 928 CAGTGTGGCTGTACGACAGTGGAGGATG-----AGGTGAG 890  
Db 498 rSerArgThrGlyGlnArgThrArgGlnArgProGlyHisProProGluArgGlySe 518  
QY 889 CAGGCCAAGAGCAGCTCTCTGGTGGTGGCCAGGT-----A 854  
Db 518 r-GlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgThrArg 538  
QY 853 GGGGGCGGACAGCTGTGTCCAGTCAATGGCAGCA-----GGAGGTAGCCAGGCA 800  
Db 538 lnArgProGlyHisProProGluArgGlySerGlyProArgGlyThrArgProProA 558  
QY 799 GCGCCCAAGACTGATCATGAAGGCATAGACAGTAGGCTGGCGCAGTGGTCCGGTC 740  
Db 558 laProSerSerArgThrGlyGlnArgThrArgGlnArgProGlyHisProProGluA 578  
QY 739 CCGAAGAGTCAAG-----GCAGGCGCTCCAGTGGAGTGAAGCAGCACACCTGGCCAGAA 583  
Db 578 rglSerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgT 598  
QY 682 GTCAGCAGCGCCAGCCAGGATGAGCAGTCCAGGCGCTCCAGGGCTGGATCCGGCA 623  
Db 598 hr-----ArgGlnArgProGlyHisProProGluArgGlySerGlyProArgGlyT 616  
QY 622 CAGCAGCGCTGTAGCCAGCGCGCTGGGATGAGAAAGAGCTCAGCAGGATGCCAA 563  
Db 616 hr-----ArgProAlaProSerSerA 624  
QY 562 GGACAGTGGCCAGATGAAGCGCGCGGCGCATAGCTCCAGCGCAGTGGTCACTGCG 503  
Db 624 rglThrGlyGlnArgThrArgGlnArgProGlyHisProProGluArgGly-----641  
QY 502 TGAGCTTAGGAGCGGACACAGACAGCAGCAGTGGACCAATGCCAGCAGCAGCATGGT 443  
Db 642 -----SerGlyProArgGlyThrArgProProAlaProSerS 654  
QY 442 CATGAAT-----435  
Db 654 erArgThrGlyGlnArgThrArgGlnArgProGlyHisProProGluArgGlySerG 674  
QY 434 -----TCTCTCTACCCCGCTCCAGCAGCAGAGCGGCGCAGTGTGCTGCTGC 383  
Db 674 lyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgThrArgGlnA 694  
QY 382 GGCACAAACACACCTCCAGCGCAAGAGTTAGCAGTTGACCA-----342  
Db 694 rglProGlyHisProProGluArgGlySerGly--ProArgGlyThrArgProProA 713  
QY 341 -----GCAAGAGTGGCTTTCGGTCCCGCAGCAGCGGCTCAGCCAGCAGCTCG 290  
Db 713 laProSerSerArgThrGlyGlnArg-ThrArgGlnArgProGlyHisProProGlu 732  
QY 289 GACCATAGTGGCGGCGCGG-----TAGGCTCAGGGCGGCTTCAGGC-----245  
Db 733 ArgGlySerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArg 752  
QY 244 -----ACTCCAGAACTGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 209  
Db 753 ThrArgGlnArgProGlyHisProProGlu-----ArgGlySerGlyProArg 769  
QY 208 GCTGCGGCTCTCTCT---CCTTCTGCGCG-----CAACTGCTTAGAATCAGCAGGC 158  
Db 770 GlyThrArgProProAlaProSerSerArgThrGlyGlnArgThrArgGlnArgProGly 789  
QY 157 GCCCATTTCTCCAGCCCTTGTGCTGGCGGTCCA---GCTTCTCAGGCCCTGCTCAACACC 101  
Db 157 -----GCTTCTCAGGCCCTGCTCAACACC 101







```

Db 405 PheSerIleLeuGlyIleProLeuAlaIleThrThrSerValProTyAlaLeuIleSer 424
QY 1721 GCCAGGTGGTTCG---GCCGGGGCATCTGGCTGGACCTGCCATCTGGATAGTGC 1777
Db 425 ThrArgIleGluSerLeuGlyGlnGlyLeuSerMetGlyValLeuAsnLeuAla 444
QY 1778 TTCCTGCTGCCAGGTGGCCCATCTGTTATGGCTCCCATGTGCCAGTCTC----- 1831
Db 445 IleValIleProGlnValIleValSerLeuGlySerGlyProTrpaspGlnLeuPheGly 464
QY 1832 -----ACCCAGTCTGTCACCTGCTATATGATGCTGCCAGCCCTGGCTGCTGTC 1882
Db 465 GlyGlyAsnSerProSerLeuAlaValAlaAlaValAlaAlaPheAlaSerGlyLeuVal 484
QY 1883 GCCATTATCTTGTCTACAGGTAGTATTGGACAGAGC 1921
Db 485 AlaIleLeuAlaIleProArgSerSerAlaAspLysSer 497
RESULT 16
O65803 PRELIMINARY; PRT; 515 AA.
AC O65803;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Sucrose/H+ symporter.
GN SUT2.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NAMTALISE; TISSUE=ROOT;
RX MEDLINE=99063785; PubMed=9847123;
RA Shaky R., Sturm A.;
RT "Characterization of source- and sink-specific sucrose/H+ symporters
from carrot.";
RL Plant Physiol. 118:1473-1480(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; Y16768; CAA76369.1; -
DR InterPro; IPR003662; sub-transporter.
DR Pfam; PF00083; sugar tr. 1.
DR TIGRfams; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 515 AA; 54424 MW; 897B90657C9E243C CRC64;

Alignment Scores:
Pred. No.: 2,53e-14 Length: 515
Score: 347.50 Matches: 131
Percent Similarity: 41.78% Conservative: 80
Best Local Similarity: 25.94% Mismatches: 215
Query Match: 5.41% Indels: 79
DB: 10 Gaps: 17

US-09-759-143-110 (1-3410) x O65803 (1-515)
QY 332 CAGCTCTTGGTGGTCAACCTGCTAACCTTTGGCTGGAGGTGTTTGGCGGAGGCATC 391
Db 34 LysLeuValLeuValAlaAlaIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu 53
QY 392 ACCTATGTCGCCCTGCTGCTGGAAGTGGGGGTAGAGAGAGTTCATGACCATGGTG 451
Db 54 SerLeuLeuThrProTyAlaGlnLeuLeuGlyIleProHisIleTrpAlaAlaTyrlle 73
QY 452 CTGGGCATTGTCAGTGGCTGGCTGCTGCTGCCGCTCTCTAGCTCAGCCAGTGCAC 511
Db 74 TrpLeuGlyGlyProIleSerGlyMetLeuValGlnProIleValGlyTyTrSerAsp 93
QY 512 CACTGGCTGACGCTATGCGCGCGCGGCTTCATCTGGGGGCACTGCTCTGGGCATC 571

```

```

Db 94 HisCysGlnSerSerPheGlyArgArgProPheIleAlaSerGlyAlaGlyCysVal 113
QY 572 CTGCTGAGCCTCTTCTCATCCCAAGGGCGGCTGGCTAGCAGGGCTG----- 619
Db 114 AlaIleSerValIleLeuIle-----GlyPheAlaAlaAspIleSerTyrlsAla 130
QY 620 -----CTGTGCCCGGATCCAGCCCTGGAGCTGGCCTGCTCATCTCTGGCGGTG 670
Db 131 GlyAspAspMetSerIleThrLeuLysProArgAlaValThrValPheValIleGlyPhe 150
QY 671 GGGCTCTGGACTTGTGGCCAGGCTGCTTCACTTCCACTGAGGCCCTGCTCTCTGAC 730
Db 151 TrpIleLeuAspAlaAlaAsnMetLeuGlnGlyProCysArgAlaLeuAlaAsp 170
QY 731 CTCCTTCCGG---GACCCGGACCCACTGTCGCCAGGCTACTGCTATCCCTTCATGATC 787
Db 171 LeuCysSerGlyAspThrArgMetArgSerAlaAsnAlaPheTyrlSerPheMet 190
QY 788 AGTCTTGGGGGCTGCTGGGCTAC-----CTCCTG 817
Db 191 AlaValGlyAsnIleLeuGlyTyAlaAlaGlySerTyrlAsnAsnLeuTyrlLysLeuPhe 210
QY 818 CCTGCCATTGACTGGGACACACAGTGCCTGGCCCTTACTGCGGCCACCCAGGAGTGC 877
Db 211 ProPheSer-----LysThrHisAlaCysAspLeuTyrlCysAlaAsnLeuLysSerCys 228
QY 878 CTCCTTTGGCTGCTCACTCCCTCATCTTCTTCCCTGCTAGCAGCCACACACTGCTGGTGC 937
Db 229 PheIleIleSerIleAlaLeuLeuIleIleThrValValAlaLeuSerValValArg 248
QY 938 GAGGAGGACCGCTGGGCCCCCAGCAGCAGGAGGCTGCTGCGGCCCTCTCTTGTGCG 997
Db 249 GluAsnSer-----GlyProAspAspAlaAspAlaAlaGluLupPro----- 263
QY 998 CCCCACTGCTGCTCATGCGCGGCGCTGCTTCCGGAACCTGGCGGCCCTGCTTCCC 1057
Db 264 -----ProSerGlyLysIleProVal-----PheGlyLeuLeuLeuGly 277
QY 1058 CGGCTGCACAGCTGTGCTGCCGATGCGCCGACCTGCGCGGCTCTTCTGGTGGCTGAG 1117
Db 278 AlaLeuLysAspLeu-----ProArgProMetLeuLeuLeuIleValThr 293
QY 1118 CTGTGAGCTGGATGACATCATGACCTTCACGCTCTTTTACACGATTTGTTGGCGGAG 1177
Db 294 CysLeuAsnTrpIleAlaTrpPheProPheIleLeuPheAspThrAspTrpMetGlyArg 313
QY 1178 GGCTCTACAGGCGCTGCCAGCTGAGCGCGGACCGCCGAGCCGAGACACTATGAT 1237
Db 314 GluIleTyrlGlyThr-----AlaGlyGlnGly-----LysLeuTyrlAsp 327
QY 1238 GAAGCGTTTCGGATGGGACGCTGGGCTGTTCTCTCAGTGGCCCATCTCCCTGCTTTC 1297
Db 328 GlnGlyValArgAlaGlyAlaLeuGlyLeuLeuLeuAsnSerValValLeuGlyLeuThr 347
QY 1298 TCTCTGCTATGACCGGCTGGTGGCAGGATTTGGCAGCTCGACATCTATTATTGGCCAGT 1357
Db 348 SerIleAlaValGluTyrlLeuValArgGlyValGlyValLysIle-----LeuTrpGly 366
QY 1358 GTGGCAGCTTCTCCTGCTGGCTGCCGCTGCCACATGCTG----- 1396
Db 367 PheValAsnPheIleLeuAlaIleGlyLeuValMetThrValValSerLysValAla 386
QY 1397 -----TCCACACAGTGTG-----GCCGTGGTGACA 1420
Db 387 GlnHisGlnArgGluHisSerAlaAsnGlyGlnLeuLeuProProSerAlaGlyValLys 406
QY 1421 GCTTACGCC-----GCCCTACCGGGTTCCACCTTCTTCACCCCTCGCATCTGCTG 1468
Db 407 AlaGlyAlaLeuSerLeuPheSerIleLeuGlyIleProLeuSerIleThrTyrlSerIle 426
QY 1469 CCCTACACTGCGCTCCCTCTTACCACCGGAGACAGGCTGCTCTCCGCCCAATACCGA 1528
Db 427 ProPheAlaLeuAlaSerIleTyrlSerSerGlySerGlyAlaGlyGlnGlyLeuSerLeu 446

```





```

Db 670 eProGlyGluArgGlyValGlnGlyProProGlyProAlaGlyProArgGlyAsnAsnG1 690
Qy 654 AGTGCCA-----GCTCCAGGGCGCTGGATCCGGG 625
Db 690 yAlaProGlyAsnAspGlyAlaLysGlyAspThrGlyAlaProGlyAla-----ProG1 708
Qy 624 CACAGACGCCCTGTAGCCAGCGCGGCGCTGGATGAGAAAGAGGCTCAGCAGGATGCC 565
Db 708 ySerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuPr 728
Qy 564 AAGGACAGCTCCAGATGAAGCGCGCGGCGCCATAGCGTCCAGCGCAGTGTCATG 505
Db 728 oGly-----ProLysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly----- 744
Qy 504 GGTAGCCTAGGAGCGGGACACAGACCGCCAGCTGGACCAATGCCAGCACCACATG 445
Db 745 -----SerProGlyLysAspGlyValArgGlyLeu-- 754
Qy 444 GTCATGAAGTCTCTCTCTA---CCCCACTTCCAGCAGCAGAGCGCGGCACATAGTGATG 388
Db 755 -----ThrGlyProIleGlyProGlyProAlaGlyAlaProGlyAspLysGlyG1 772
Qy 387 CTGCGCGCCCAACACACACT-----CCAGGCCAAAGG-- 357
Db 772 uAlaGlyProSerGlyProAlaGlyProThrGlyAlaArgGlyAlaProGlyAspArg1 792
Qy 356 -TTAGCAGGTTGACCAAGAGAGCTGGCTTTCGGT---GCCGAGCAGGCGGCTCACC 301
Db 792 yGluAlaGlyProProGlyProAlaGlyPheAlaGlyProProGlyAlaAspGlyGlnPr 812
Qy 300 CACAGCCTTGACCACTAGTGGG-----CCAGGGGGTAGGGCT 262
Db 812 oGlyAlaLysGly-GluProGlyAspThrGlyValLysGlyAspAlaGlyProProGlyP 832
Qy 261 CAGGGGCGCTCAGGCACTCCAGAACTCGCTCGCTCGCTCCCA----- 212
Db 832 roAlaGlyProAlaGlyProProGlyProIleGlyAsnValGlyAlaProGlyProLysG 852
Qy 211 -----GAAGCTGGGCTCTCTCTCTGTCGCCCAACTGCTAGGAATCAGCAGCGC 157
Db 852 lySerArgGlyAlaAlaGlyProPro-GlyAlaThr-----GlyPheProGlyAla 868
Qy 156 CCCATTTCGCCAGCCCTTGTGCGCGTCCAGCTTCTCAGCCCACTGCTCACACCTCT 97
Db 869 AlaGlyArgValGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProGly 888
Qy 96 GCTGGGGCAGCTCAGT-----GGGACACGCTCTCATCTCAGCTCAGTCCCT 52
Db 889 ProValGlyLysGlyGlyGlyGlyProArgGlyGluThrGlyProAlaGlyArgPro 908
Qy 51 GGGCGA-----GGGCGCGGCTGCTCAGCGGAGCC 22
Db 909 GlyGluValGlyProProGlyProGlyProAlaGlyGluLysGlySerProGlyAla 928

```

RESULT 14  
O93251

```

ID O93251 PRELIMINARY; PRT; 1445 AA.
AC 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DT Alpha 1 type I collagen.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99294154; PubMed=10367734;
RA Asahina K., Utoh R., Obara M., Yoshizato K.;
RT "Cell-type specific and thyroid hormone-dependent expression of genes
RT of  $\alpha 1(I)$  and  $\alpha 2(I)$  collagen in intestine during
RT amphibian metamorphosis."

```

```

RL Matrix Biol. 18:89-103(1999).
DR EMBL; AB015440; BAA29028.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF01410; COLFI; 1.
DR ProDom; PD000007; Collagen; 3.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; VWFC; 1.
DR PROSITE; PS01208; VWFC; 1.
KW Collagen.
SQ SEQUENCE 1445 AA; 137252 MW; F59BB550C9873F04 CRC64;

```

# Alignment Scores:

Pred. No.:	1-06e-14	Length:	1445
Score:	354.50	Matches:	256
Percent Similarity:	31.75%	Conservative:	51
Best Local Similarity:	26.47%	Mismatches:	307
Query Match:	5.69%	Indels:	353
DB:	13	Gaps:	58

US-09-759-143-110 (1-3410) x O93251 (1-1445)

```

Qy 2588 TGTCTCTGTGTGCAACAGAGACCAACAGCCACATCTGTATAAAGGTAAAGGGG 2529
Db 33 CysValValAspGly-----ArgThrTyrAsnAspLys----- 43
Qy 2528 GGTGGATCACCACAAACACAGTGTGTGGCTGAGGGGACCTGTTCTTGTGTGTGCC 2469
Db 44 ---AspValTrpLysProGluAlaCysGlnIle-----CysVal-CysAs 57
Qy 2468 CTCAGGACTCTTCCCTACAAATAGTCATATGTTCAATCCCATGGAGGAGTGTTCAT 2409
Db 57 pGluGlyThrIleLeuCysAspGluValIleCysGluAsp---IleGlyAspCysProAs 76
Qy 2408 CCTAGAACTCCCATGCAAGAGCTTACATTAAACGAAGCTCAGGTTAAGGGCTTAGAGA 2349
Db 76 nProGluIleProMetGlyGlyLucys-----CysProValCysGlyGluGlyG1 92
Qy 2348 TGGGAACACAGGCTGACTGAGTTTATTCAGCTCCCAAAACCCCTCTCTAGTGTGTCTCA 2289
Db 92 nTyrGlnThrGlySer---ValValGluGlyProLysGlyGluThr----- 106
Qy 2288 ACTAGGAGGCTAGTGTGTTAAACCTGAGCCTGGGTGTAATCCACCTGCAGAGTCCCGCATTC 2229
Db 107 -----GlyProArgGlyGluArgGlyProPr 115
Qy 2228 CAGTGCATGAGCCCTTCTGCGCTCCCTGTATAGTCCAGACTCAAAACCCCTTGGGAAGG 2169
Db 115 oGlyAlaProGlyArgAspGlyIleProGly----- 125
Qy 2168 CCTCCAGTCAGCAGCCCTAGAGACTGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2109
Db 126 ---GlnProGlyIlePro-----GlyProProGlyProProG1 137
Qy 2108 TGTGCAGCTAGCCACCTCAGCAGCAGAGGTGCGGCTAACAGAGGAGGAGGAGGAGGAGG 2049
Db 137 y-----ProAlaGlyLeuGlyGlyAsn-----PheAlaPr 147
Qy 2048 CAACAGAACTGGCGGCGGAGCCCGGAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1989
Db 147 oGlnMetSerTyrGlyTyrAspGlyLysSerAlaGlyIleSerMetProGlyProMetG1 167
Qy 1988 ACCCAGTGAGCAGGCGCCCTCACCCCAATGTGTGGAAGTGTTCAGCTGAGTATTGG 1929
Db 167 yProMetGlyProArgGlyProProGlyProSerGlySer----- 180
Qy 1928 CCAAGTCGCTCTTGTCAAATACTACTACCTGTGTGTAGCAAAAGTAATGGCAGCAGCCAGCC 1869
Db 181 -----ProGly--ProGln---Gly 185

```

QY 2501 -----GGGCTGAGGGACCTGGT 2484  
Db 85 PheCysProGluGluTyrValSerProAspAlaGluValIleGlyValGluGlyPro--- 103  
QY 2483 TCTTGTGTGTTGGCCCTCAGGACTCTTCCCTACAATAAGTCAATATGTTCAATCCCAT 2424  
Db 104 -----LysGlyAsp 106  
QY 2423 GGAGGAGTGTTCATCTCAGAACTCCCATCAAGAGCTACATTAAAGAGTGCAGGT 2364  
Db 107 ProGlyProGlnGlyProArgGly-ProValGlyPro----- 118  
QY 2363 TAAGGGCTTAGAGTAGGAAACAGGTGACTGAGTTATTTCAGCTCCCAAAACCCCT-- 2306  
Db 119 -ProGly--GlnAspGlyIleProGlyGlnProGlyLeuProGlyProPro 137  
QY 2305 ----TCTCTAGTGTCTCAACTAGGAGGTAGCTGTTAAC----- 2267  
Db 137 oGlyProProGlyProLeuGlyLeuGlyGlyAsnPheAlaSerGlnMetSertyrGlyTy 157  
QY 2266 ----CTGAGCCTGGGTAAATCCACCTGCCAGAGTCCCGCATTCAG 2226  
Db 157 AspGlySerAlaGlyValSerValProGlyProMetGlyProser----- 173  
QY 2225 TGCATGGAGCCCTTCTGGCCCTCCCTGTATAGTCCAGACTGAACCCCTTGGAGGCCT 2166  
Db 174 -----GlyProArgGlyLeuPro----- 179  
QY 2165 CCAGTCAGGACCCCTAGAGACTGGGAGAGAGAGAGGAGCGCCAGCCCGCCAGCTGT 2106  
Db 180 -----GlyProPro-----GlyAlaProGlyPro----- 187  
QY 2105 GCAGTCAGCACCTCAGCAGCAGAGGTGGCAGCAGAGCACATTACTTTGSCACAA 2046  
Db 188 -----GlnGlyPheGlnGlyProProGlyGluProGlyGluProGlyGlyse 203  
QY 2045 CAGAACTGGCGGCCAGCCCGCCATGGGGCTTAACGAGCGGGAGCTGGGA-- 1988  
Db 203 rGlyProMetGlyProProGlyProProGlyProProGlyLysAsnGlyAspAspGlyL 223  
QY 1987 -----CCAGTCAGGAGGCCCTCCA----- 1967  
Db 223 uAlaGlyLysProGlyArgProGlyGluArgGlyProProGlyProGlnGlyAlaArgL 243  
QY 1966 ----CCCCAATGTCTGGAAGTTTCTACGCTGATATTGGCCAAAGTCGCTCTGTCAA 1911  
Db 243 yLeuProGlyThrAlaGly-----LeuProGlyMetLysGlyHlsArgL 258  
QY 1910 ATACTACTGTGTAGCAAGTAATGGCGACACCCAGCCCTGGCGCAGACACCATAT 1851  
Db 258 yPheSerGlyLeuAspGlyAlaLysGlyAspThrGlyProAlaGlyProLysGlyGluPr 278  
QY 1850 AGGCAGTCA----CAGACTGGCTGAGTGGACATGGAGCCCATAAACAGGATG----- 1800  
Db 278 oGlySerProGlyGluAsnGlyThrProGlyGlnMetGlyPro---ArgGlyLeuProG 297  
QY 1799 -----GGCCACCTGTCACAGCAGGAGGACCTATCCAGATGGCGAG 1756  
Db 297 yGluArgGlyArgProGlyProProGlyThrAlaGlyAlaArgGly-AsnAspGlyAlaV 317  
QY 1755 TCCAGGACAGATGCCCGGCCCGGAAACCCCTGGCTGGCTCAGCCACCCACACCA 1696  
Db 317 alGlyAlaAlaGlyPro-ProGlyProThr---GlyProThrGlyProProGlyPhePro 335  
QY 1695 CGTACGAGACATCACAGGACAGGCC--CCGAGAGCGGGGTGGAGT----- 1648  
Db 336 GlyAlaAlaGlyAlaLysGlyGluAlaGlyProGlnGlyAlaArgGlySerGluGlyPro 355  
QY 1647 ---GGGAGCAGG-----CCACTGCCTCCAGCACCCAGCTGTCCATTAGGG--AAGGGA 1600  
Db 356 GlnGlyValArgGlyGluProGlyProProGlyProAlaGlyAlaAlaGlyProAlaGly 375  
QY 1599 GCTCAGGCTTAGGG-----CCTGGCAGGAAGCTGGTCAATCAGGCTGTCTCACTACTA 1546

Db 376 AsnProGlyAlaAspGlyGlnProGlyAlaLys-GlyAlaAsnGlyAlaPro----- 392  
QY 1545 GCACCTCCAGCTGTCCTCGTATTGGCGCAGGAAACACCTGCTTCTCCCGGTGGTACAGG 1486  
Db 393 -----GlyIle---AlaGlyAlaProGlyPheProGlyAlaArg 405  
QY 1485 GAGGCCAGTGTGTAGGCGAGGATCTCAGGGCTGAGAAGGTGAACCCGCTGAGGCGGCT 1426  
Db 405 y-----ProSerGlyPro-GlnGlyProSerGlyAlaProGlyProLysGlyT 421  
QY 1425 GAAGCTGTCAACCGCCACACTGTGGGACAGGATGTGGCCAGCCACAGGAA 1366  
Db 421 hrSerGlyGluProGly---AlaProGlyAsnLysGlyAspThrGlyAla--LysGlyG 439  
QY 1365 GCTGCCACACTGCCCAATAGACTGCTCAGTCCGAGTCCGAATCGCTGCACAGCCGT----- 1311  
Db 439 u-----ProGlyProAlaGlyValG 446  
QY 1310 -----CCATGACAGAGAGAGACAGGAGATGGCGACTGCAGGAACA----- 1266  
Db 446 nGlyProProGlyProAlaGlyGluGlyLysArgGlyAlaArgGlyGluProGlyPr 466  
QY 1265 -----GCCCGAGCTGCCATCCGACGC-----CT 1240  
Db 466 oSerGlyLeuProGlyProProGlyGluArgGlyGlyProGlySerArgGlyPheProG 486  
QY 1239 TCATCATAGTGTCTCCGGCCCTCGTGGCGGTCTAGCTCTGGGCACCCCTGGTACAGC 1180  
Db 486 yAlaAspGlyValAlaGlyProLysGlyProSerGlyGluArgGlySerProGlyProAl 506  
QY 1179 CCCTCGCCACAGAAATCCGTGTAAACAGCGCTCAAGTCAAGTCCATGCCATCCAGCTGCAC 1120  
Db 506 aGlyProLysGlySerPro----- 512  
QY 1119 AGCTACGCCACGAAGAGCGCGCGGATGCGGGCATGCGCAGCAGCTGGTGCAGC 1060  
Db 513 -----GlyGluAlaGlyArgPr 518  
QY 1059 CGGGAGAGCAGGCGCCAGGT-----TCCGGAAGCCAGGGCGCCCGCAT 1012  
Db 518 oGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGlySerProGlySerProGlyPr 538  
QY 1011 GGACAGCAGTGGCGGCACAGAGGCGCGCAGACGCCCTTCTGCTGGCTGGTGGGCGCC 952  
Db 538 oAsp-----GlyLysThrGlyProProGlyProAlaGlyGlnAspGlyArgProGlyPr 556  
QY 951 AGCGCTGCCT-----CCTCAGCCACCCAGC 928  
Db 556 oAlaGlyProProGlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyProLysG 576  
QY 927 AGTGTGGCTCTACGCGAGGTGAGGAAGATGAGGTGAGCGGCAAGAGGCAAGGCACTCT-- 870  
Db 576 yThrAla-----GlyGluProGlyLysAlaGlyGluArgGlyValProG 591  
QY 869 -----CCTGGGTGCC 859  
Db 591 yProProGlyAlaValGlyProAlaGlyLysAspGlyGluAlaGlyAlaGlnGlyAlaPr 611  
QY 858 AGGT---AGGGGCCAGGCGACTGGTGTCCAGCTCAATGGCAGGCAAGG----- 813  
Db 611 oGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyProAlaGlySerProGlyPh 631  
QY 812 -----GGTAGCCCGCAGCGCCCAAGACTGATCATGAAGCATAGACAGAG 766  
Db 631 eGlnGlyLeuProGlyProAlaGlyProProGlyGluAlaGlyLysProGlyGluGlnG 651  
QY 765 TAGGCTGGCGCAGCTGGT---CCGGGTCCC-----GGAAGAGGTTCAGAGACGAGGCC 715  
Db 651 yValProGlyAspLeuGlyAlaProGlyProSerGlyAlaArgGlyGluArg---GlyPh 670  
QY 714 TCCAGTGGAGTGAAGCACACCTCGCCACAGAGTCCAGAGCGCCCGCCAGGATGAGC 655

```
QY 1149 GTGAAGGTCATGAGTGCCA-----TCCAGCTGCACAGCTCAGCCACGCAAG 1105
Db      |||||
QY 296 -----ValProGlyGlyMetGlyGlyThrGlyGlyProArgse 311
Db      |||||
QY 1104 ACCCGCGCAGGTCGGGGCA-----TGGCGCAGCAGCTGGTGAGCCGGG----- 1056
Db      |||||
QY 311 rGluglyAlaGlyCysProAlaArgGlyAlaGlyProAlaTrpGlyAlaGlyProPr 331
Db      |||||
QY 1055 -----GAAGCAGGCGCCAGGTTCGGAAAGCCCAAGC-----GG 1021
Db      |||||
QY 331 oArgArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySerGlyLeuProGl 351
Db      |||||
QY 1020 GCCCGCATGACAGCAGTGGGCGCACAAAGAGGGCGCCGACAGCC----- 975
Db      |||||
QY 351 yProGlyCysGlySerProGlyLeuGlySerArgSerProAlaGlyTrpGlyGluPr 371
Db      |||||
QY 974 -----CTTCTGCTGCTGGTGGGCGCCAGGCTGCCT-- 942
Db      |||||
QY 371 oGlyGlyArgGlyProGlyAlaArgGluArgAlaArgProGlyValArgValProAr 391
Db      |||||
QY 941 -----CCTCAGCCAGCAGCAGTGTGCTAGCGCAGGTGAGGAAGATGAGGGTG 892
Db      |||||
QY 391 gProGlyGluProValProProGly-----GlyMe 401
Db      |||||
QY 891 AGCAGGCCAAAGAGCAGCTCCTCTCT-----GGTGCCCGCAGTAGGGGCC 847
Db      |||||
QY 401 tGlyGlyThrGlyGlyThrGlyProArgSerGluGlyAlaGlyCysProAlaArgGlyAl 421
Db      |||||
QY 846 AGGCGACTGGTGTCCAGTCAATGCCAGGCAGGA-----GG 811
Db      |||||
QY 421 aGly-----ProProAlaTrpGlyAlaGlyProProArgArgAspGlyGlyAsnAr 438
Db      |||||
QY 810 TAGCCAGCAGCAGCCCGCAGTGTATGATGAAGGATAGACAGAGTAGGCTGCGCAG 751
Db      |||||
QY 438 gGlyAspGlyAlaProGluArgGlySer-----GlyLeuPr 450
Db      |||||
QY 750 TGGTCCGGGT-----CCCGGAAGAGGTGACAGACAGCGGCTCCAGTGGAGTGAAG 700
Db      |||||
QY 450 oGlyProGlyCysGlySerProGlyLeuGlySerArgSer---ProProAlaGlyTrpGl 469
Db      |||||
QY 699 CACACTGCGCCACAGTCCAGCA-----GCCCGCAGCCAGCAGTACGAGT 652
Db      |||||
QY 469 yGluProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgVa 489
Db      |||||
QY 651 GCCAGCTCCAGGGCGCTGGATCCG-----GCCACAGCAGCCCTGCTAGC 607
Db      |||||
QY 489 lProArgProGlyGluProValProProGlyGlyMetGlyGlyThrGlyGly----- 506
Db      |||||
QY 606 CAGCCGCGCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCAAGGACAGTGCACAGATG 547
Db      |||||
QY 507 -----ThrGlyProArgSerGluGlyAlaGlyCysProAlaArgGlyAlaGlyPr 523
Db      |||||
QY 546 AAGGCGCGGCGCGCCATAGGCTCCAGCCAGTGTCTACCTGGCTAGCAGCGGG 487
Db      |||||
QY 523 oProAlaTrpGlyAlaGlyProProArgArgAspGlyGlyAsnArgGlyAspGlyAlaPr 543
Db      |||||
QY 486 ACACAGA-----CCAGGCGCCAGCAGCTGGACCAATGCCAGCAGCAGTGTCTATG 439
Db      |||||
QY 543 oGluArgGlySerGlyLeuProGlyProGly-----CysGlySerPro---GlyLe 559
Db      |||||
QY 438 AACTTCTCTTACCCCGCAGTCCAGCAGCAGGAGCGGCGGCATAGGTGATGCCGCGGCC 379
Db      |||||
QY 559 uGlySerArgSerProProAlaGlyTrpGlyGluProGlyGlyArg----- 574
Db      |||||
QY 378 AAACACACTCCAGCGCCAAAGTTAGCAGGTTGACCAAGCAGCAGTGGCTTTCCGGTGC 319
Db      |||||
QY 575 -----GlyProGlyAlaArgGluArgAla-----AlaArgProGlyValArg-ValP 590
Db      |||||
QY 318 CGCAGCAGCGCGCTCAGCCACAGCAGCTCTGGACCACATAGTGGCGCAGCGGGTAGGCTCAG 259
Db      |||||
QY 590 ro---ArgProGlyGluProValPro-----ProGly-----GlyMetG 602
Db      |||||
```

```
QY 258 GGGCGCGTTTTCAGGCACTCCAGCAAGCTCTTCTGGCTCTCTCCAGAAAGCTGGCGCT 199
Db      |||||
QY 602 LygGlyThrGlyGlyThrGlyProArgSerGluGlyAlaGlyCysProAlaArgGlyAlaG 622
Db      |||||
QY 198 CTCTCTCTTGTCTGCCCAACTGCCCTAGGAATCAGCAGCGGCGCCATTTCTGCCAGCCCT 139
Db      |||||
QY 622 LyProProAlaTrpGlyAlaGlyProProArgArgAspGlyGlyAsn-----ArgGlyA 640
Db      |||||
QY 138 TTGCTGCTC-----GGTCCAGCTTCTCAGCCCATGCTCAACA 103
Db      |||||
QY 640 spGlyAlaProGluArgGlySerGlyLeuProGlyProGly----- 653
Db      |||||
QY 102 CTGCTGCTGTGGGCACTCAGTGGGACACGCTCTCA 65
Db      |||||
QY 654 -----CysGlySerProGlyLeuGlySerGlnSer 663
Db      |||||

RESULT 13
Q63079          PRELIMINARY;          PRT; 1453 AA.
ID   Q63079;
AC   Q63079;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DI   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Collagen alpha1 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
RX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE OF 1-1092 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=TOOTH;
RA   Brandstetter C., Lundmark C., Christersson C., Hammarstrom L., Wurtz T.;
RT   "Expression of Collagen alpha1(i) mRNA variants during Tooth and Bone
RL   J. Dent. Res. 0:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=TOOTH;
RA   Wurtz T.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR   EMBL: Z78279; CAB01633.1; -.
DR   InterPro: IPR000087; Collagen.
DR   InterPro: IPR000885; Fib_Collagen_C.
DR   InterPro: IPR001007; VWF_C.
DR   Pfam: PF01410; COLFI; 1.
DR   Pfam: PF01391; Collagen; 18.
DR   ProDom: PD000007; Collagen; 3.
DR   ProDom: PD002078; Fib_collagen_C; 1.
DR   SMART: SM00038; COLFI; 1.
DR   SMART: SM00214; VWF; 1.
DR   PROSITE: PS01208; VWF; 1.
KW   Collagen.
FT   NON_TER
FT   NON_TER
SQ   SEQUENCE 1453 AA; 137887 MW; E6896BDC19A4A1D8 CRC64;
```

```
Alignment Scores:
Pred. No.: 7,32e-15 Length: 1453
Score: 357.00 Matches: 283
Percent Similarity: 31.84% Conservative: 50
Best Local Similarity: 27.06% Mismatches: 362
Query Match: 5.73% Indels: 352
DB: 11 Gaps: 59

US-09-759-143-110 (1-3410) x Q63079 (1-1453)
```

```
QY 2594 TGCTGTGTGT-----CTCTGTGTGTGGC-----AACACAGGACCAAC 2559
Db      |||||
QY 52 CysileCysHisAsnGlyThrAlaValCysAspGlyValLeuCysLysGluAspLeuAsp 71
Db      |||||
QY 2558 AGGCCACATCTGATAAAGTAAGGGGGTGGATCAGCAAAAGACAGTGTCTGT--- 2502
Db      |||||
QY 72 CysProAsnProGlnLysArg---GluGly-----GluCysCysPro 84
Db      |||||
```





```

QY 166 CAGCCAGGCGCCATTTCTGCGAGCCTTGGTGGTCCAGCTTCTCAGCCCATGCTC 107
Db |||||||
QY 961 SerProGlyAla-----ValGlyProAlaGlyAsnPro----- 971
Db |||||||
QY 106 AACACCTGCTGCTGGGGACCTCACTAGTGGGACACCTCTCATCATCACTAGCTGCTG 49
Db |||||||
QY 972 Gly-ProAlaGlyGluAsnGlyMetProGlySerAspGlyAsnAspGlyAlaProGlyPr 991
Db |||||||
QY 48 -----CGAGCGCGGGCTGTCCACCCGAGGCGGCGCTGCGAGGCTGTTCC 1
Db |||||||
QY 991 oGlnGlySerArgGlyGluGlyGlyAspThrGlyAlaSerGlyAlaAsnGlySerPro 1010
Db |||||||

RESULT 11
Q905K9 PRELIMINARY; PRT; 608 AA.
AC Q905K9;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
DE NTR.
OS Herpesvirus papio.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BARCOON LYMPHOCRYPTOVIRUS BA65;
RA Zong J.-C., Ryan J., Ling P.D., Loeb D.D., Pagano J.S., Hayward S.D.,
RA Hayward G.S.
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF200364; AAF23950.1; -.
SQ SEQUENCE 608 AA; 60458 MW; 4BF82ACB0A029627 CRC64;

Alignment Scores:
Pred. No.: 3,51e-15 Length: 608
Score: 361.00 Matches: 203
Percent Similarity: 35.05% Conservative: 28
Best Local Similarity: 30.80% Mismatches: 263
Query Match: 5.80% Indels: 165
DB: 12 Gaps: 36

US-09-759-143-110 (1-3410) x Q905K9 (1-608)
QY 1805 GGGATGGGCGCCACTTGGGACA-----GCAGAGGCGACTATCCAGGATGG 1761
Db |||||
QY 11 GlyProGlyArgProGlyThrPheProGlyArgGluAlaGlyArgGluAlaGlyArgThr 30
QY 1760 CGAGTCCAGGCAGATCCCGCGCGCCGGAACCCACCTGGCTCGGTGGGCTCACCACCA 1701
Db |||||||
QY 31 TrpGlyPro-----CysProAlaProSerProProAlaPheGlnValHisGlyPro 48
QY 1700 CCACAGCTACGAGACATCAGCAGGAGCGCGCGAGCGCGGT-----G 1653
Db |||||
QY 49 ArgSerProArgThrGluAlaArg-ArgGlySerAlaGlnArg-GlyHisProProG 68
Db |||||
QY 1652 GAGTGGGAGGAGCGCCACTGCTCCAGCACCACCTGCTCATAGGGAAGGAGCTCCAG 1593
Db |||||
QY 68 lYalaGlyGlnArgProSerGlyProThrGlyHisProAlaAlaProGlyAlaProG 88
QY 1592 GC---TTAGGCGCTGGCA-----GGAAGCTGGTCACTCAGGCTGTCTCCAC 1551
Db |||||
QY 88 lYProArgSerPro-ArgThrGluArgArgArgGlySerAlaGlnArgGlyHisProPro 107
QY 1550 TGCTAGCACCTCCAGTCCCTCGGTATTTGGCAGAGAACACCTGCTTCTCCCGGTGGT 1491
Db |||||
QY 108 ProGlyAlaGlyGlnArgProSerGlyProThrGlyHisProAlaAlaProGlyAla 127
QY 1490 AGAGGG-----AGCCAGTGTAGGCGAGATCTGCAGGGCTGAGNAGTGA 1443
Db |||||
QY 128 ProGlyProArgSerProArgThrGluArgArgGlySerAla----- 142
QY 1442 ACCCGGTGAGGCGGCTGAAGCTGTCCACCGGCCACACTGTGGGACAGCATGTGGCAC 1383

```

```

Db |||||
QY 143 -----GlnArgGlyHisProProGlyAlaGly----- 152
QY 1382 CGCAGCCACACAGGAAAGCTGCCACACTGGCCCAATAGACTCTCGAGTGGGAATCGCT 1323
Db |||||
QY 153 GlnArgProSerGlyProThrGlyHisPro-----Ala 164
QY 1322 GCACAGCGCGTCCATGACAGAGAGACAGCAGGAGATGGCCAGTGCAGAGAACAGCC 1263
Db |||||
QY 165 AlaProGlyAlaProGlyPro---ArgSerProArgThrGluArgArgGlySerAla 183
QY 1262 CGAGGCTGCCATCCGACGCGCTTCATCATAGTGTCTCGGCGCTCGGTGCCCGCTCAG 1203
Db |||||
QY 184 -GlnArgGlyHisPro-----ProProGlyAlaGlyGlnArgProSe 197
QY 1202 C-----TCTGGGCGAC-----GCCCTGTACAGCCCTCGCCACGAAATCCGTGT 1158
Db |||||
QY 197 rGlyProThrGlyHisProAlaAlaProGlyAlaProGlyProArgSerProArgThr 217
QY 1157 AAACACGCGTGAAGTTCATGCTGCTCCGGAAGCCAGCGCGCGCTGACAGACAGCGG 1098
Db |||||
QY 217 rGluArgArgGly-----SerAlaGlnArgGlyHisProProG 232
QY 1097 GCAGGT-----GCGGGCATGCGGCGACAGCAGCTGCTGCAGCC 1059
Db |||||
QY 232 yAlaGlyGlnArgProSerGlyProThrGlyHisProAlaAla-----Pr 248
QY 1058 GGGGAAGCAGCGCGCGCTCCGGAAGCCAGCGCGCGCTGACAGACAGCGAGTGGG 999
Db |||||
QY 248 oGlyAlaProGly-ProArgSerProArgThrGluArgArgGlySerAlaGlnArgG 268
QY 998 GC---GACAAAGAGGGCGGCGACAGCCCTTCTGCTGCTGGTGGGCGCCAGCGCTGCT 942
Db |||||
QY 268 lYHisProProGlyAlaGlyGlnArgProSerGlyProThrGlyHisProAlaA 288
QY 941 CCTCAGCCACAGCAGTGTGGCTGTACGAGTGAGGAGATGAGGTGACAGCGCCAA 882
Db |||||
QY 288 lAProGlyAlaProGlyProArgSerProArgThrGluArgArgGlySer-AlaGln 307
QY 881 AGAGGACATCTCTCTGGTGGTGGCGAGTGGGCGCA-----GGGCACTGG 837
Db |||||
QY 308 ArgGlyHisProProGlyAlaGlyGlnArgProSerGlyProThrGlyHis 326
QY 836 TGTCCAGCTCAATGGCAGCAGGAGTACCCAGCAGCGCCCGCCAGACTGATCATGAGG 777
Db |||||
QY 327 ---ProAla-----AlaProGlyAlaProGly---ProArgSerProArgThrGlu 341
QY 776 CATAGACAGTAGGCTGGCGACAGTGTGCGGGTCCCGGAAGAGGTACAGAGA----- 723
Db |||||
QY 342 ArgArgArgGlySerAlaGlnArgGlyHisProProGlyAlaGlyGlnArgProSer 361
QY 722 -----GCAGGGCTCCAGTGGAGTGAAGCACACCTGGCCAGAGTCCAGCAGCC 672
Db |||||
QY 362 GlyProThrGlyHisProAlaAlaProGlyAlaProGlyProArgSerProArgThr 381
QY 671 CCAGCGCCAGGATGACAGCTG-----CCAGCTCCAGGGCGCTGGGATCCGGGC 624
Db |||||
QY 382 GluArgArgGlySerAlaGlnArgGlyHisProProGlyAla----- 397
QY 623 ACAGCAGCCCTGTGATCCAGCGCGCTTGGGTGAGTGAAGAGGCTCAGCAGGATGC--- 567
Db |||||
QY 398 -----GlyGlnArgProSerGlyProThrGlyHisProAlaAlaPro 412
QY 566 -----CCAGGACAGTGGCCAGATGAAGGCGCGCGCGCGCCAT 528
Db |||||
QY 413 GlyAlaProGlyProArgSerProArgThr-----GluArgArgGlySerAlaGln 430
QY 527 AGCGTCCAGCGCAGTGGTCACTGGCTGAGCCCTAGGAGCGGCGACAGACCA---GGCCCA 471
Db |||||
QY 431 ArgGlyHisProPro-----GlyAlaGlyGlnArgProSerGlyPro 445
QY 470 GCACTGAGCAATGCCAGCAGCAGCTGTCATGAATCTCTCTACCCCTCCAGCA 411
Db |||||

```

656 hrGlyAlaGlnGlyArgArgGlyIleGlyGlyArgAlaGlyAsnTyrGlyAlaThrGlyp 676

1067 -----GTCAGCCGGGAAAGCAGAGCGCCAGGTCC-----G 1034

676 roProGlyGlnLysGlyGluMetGlyProProGlyAsnValGlyLeuGlnGlyProProG 696

1033 GAAAG-----CCAAGCGGCCCGGCATGCACAGCTGGGGGCACAAAGAGGGGCCGA 980

696 LyGluLeuGlyProSerGlyPro-----ProGlyAla-----ArgGlyProp 710

979 CAGCCCTTCGCTGCGCTCGGTGGGGCCCGCGCTGCTCTCAGCCACAGCAGCTGTGC 920

710 ro-----GlyProSerGlySerProGlyProAasp-----719

919 TGCTACGAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGCAGCT-----873

720 -----GlyProAlaGlyAlaGluGlyAspArgGlyProValGlyPrometG 735

872 -----CCTCTCGGTGCCAGGT---AGGGG-----CCAGGGCAGCT 839

735 LyProSerGlyProSerGlyMetProGlyGluArgGlyAspAsnGlyGluProGly---753

838 GGTGTCCAGTCAATGCCAGCAGGAGGTAGCCAGCAGCCGCCCAAGACTGATCATGAA 779

754 -----ProGlyAlaProGlyGlnArgGlyAasp-----763

778 GGCATAGACAGAGTAGCCCTGGGCACAGTGGTCG-----GGTCCGGGAAGAGTCA 725

764 -----ProGlyAspLeuGlyProGlnGlySerProGlySerProGlyp 778

724 GAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCGC 666

778 heAlaGlyProProGlyArgSerGlyAsnProGlyProGlnGlyGluLeuGlyProThrG 798

665 -----CCAGGATGAGCAGTCCAGCTCCAGGCTCCAGGGGCTGGGATC 629

798 LyAlaArgGlyGluThrGlyGlyProGly-----ProSerGlyProThrGlyAasp 815

628 CGGCACAGCAGCCCTCTAGCCAGCGCGCCCTGGGTGGATGGAAGAGCTCAGCAGGAT 569

815 roGlyProGlnGlyProLeuGlyAlaProGlyGlnGlnGlyGluArgGlyGluThrGly- 834

568 GCCCAAGCAGAGTCCCGAGATGAAGCGCCGCGCGCCCATAGCGTCCAGCCGAGTGT 509

835 -----ProGln-----836

508 ACTGGCTGAGCCTAGGACGGGCACACAGACAGCCAGCCAGCAGCTGGACCATGCCAG 449

837 -----GlyGlnGlyGlyProGlyPro-----IleGlySerLeuGlyAlaP 851

448 CATGGTCATGAACCTCTCTACCCCTCCAGCAGCAGAGGGCGGCACATAGGTGAT 389

851 roGlyAlaGlnGlyProProGlyProThrGlyProSerGlyAsnAlagly-----867

388 GCCTCGCGCCCAACACACCT-----CCAGGCCAAAGATTAGCAGGTTGAC 344

868 -----SerProGlyGlnProGlyAlaArgGlyGluProGlyGln-----SerGlySerP 884

343 CAGCAAGAGCTGGGCTTCCGGTGCCCGCAGCA-----GGCGGCT 305

884 roGlyGlnProGlyLeuAlaGlyArgThrGlyProSerGlyGluArgGlyAspGlyAla 904

304 CACCCACAGCCTCTGGACCATAGTGG-----CCAGCGGGTAGGCTCAGGGGCCCTTCA 248

904 snAspGlyGlnSerGlyPro-ProGlyProProGlyProAlaGlyProAlaGlyGlnSer 923

247 GGCACCTCCAGAACTGCTTGTCTCT-----GGCTCTCTCCTCAGCAAGCTCGGCTCTCT 194

924 Gly-----IleLeuGlyLeuAlaGlyGlySerGlyProArgGlyProGlyGlyPro 940

193 -----CCTTGTCTCGCGCCCAACTGCCTAGGAAT 167

941 AlaGlyProProGlyAlaAlaGlySerArgGlyProAlaGlyLysSerGlyAspArgGly 960





```

Db 278 euTyrPheAlaGluValProLeuMetAlaTyrGlnProHisHisLeuSerAspSer- 297
QY 977 CTGTGGCCCTCTCTGTGCGCCACGTGCTGTCATCGCGCGCGCTTGCTTT----- 1032
Db 298 -----AlaProLeuLeuAspAsnProGlnGlnIleGlyPheAspA 311
QY 1033 -----CCGNA- 1038
Db 311 snSerLysSerLysLeuAspMetSerAlaValAspAlaThrGlyAsnAsnProGluS 331
QY 1038 ----- 1038
Db 331 erSerTyrGluIleAsnLysAsnAlaLysHisLeuThrProIleValGlnGlnAsnG 351
QY 1039 -----CCTGGGCGCTGCTTCCCGCGGTGCGCCACGTGCTGTGCGCGC 1081
Db 351 luSerPheSerAspGlyPro-GlyAlaValLeuValAsnLeuLeuThrSerLeuArgHis 370
QY 1082 ATGCCCGCGCACCTCGCGGCTCTCTGCTGCTGAGCTGCGAGCTGGATGGCCTCATG 1141
Db 371 LeuProProAlaMetHisSerValLeuLeuValMetAlaLeuSerTrpLeuSerTrpPhe 390
QY 1142 ACCTTCACGCTTTTACAGGATTCCTGGCGAGGCGCTCTACAGGCGGTGCCGAGA 1201
Db 391 ProPhePheLeuPheAspThrAspTrpMetGlyArgGluValTyrHisGlyAspProLys 410
QY 1202 GCTGACCGCGGCGCGGCGGAGACACTATGATGAAGCGTTCGGATGGCAGCGCTG 1261
Db 411 GlyAsp-----GluSerAlaValLysAlaTyrAspAlaGlyValArgGluGlyAlaPhe 428
QY 1262 GGGCTGTCTCTGCTGCGGCTCTCTCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTG 1321
Db 429 GlyLeuLeuLeuAsnSerValAspLeuGlyIleSerPheLeuIleGluProMetCys 448
QY 1322 CAGCGATTCGGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1378
Db 449 GlnArgMetGlyAlaArgLeuValTrpAlaMetSerAsnPheIleValPheAlaCysMet 468
QY 1379 GCGGTGCGCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1399
Db 469 AlaGlyThrAlaIleIleSerLeuValSerValAsnGluTyrIleThrGluGlyIleGln 488
QY 1400 CACAGTGTG-----GCCGTGCTGACAGCTTCAGCC-----GCCCTCACC 1438
Db 489 HisAlaIleGlyGluAsnArgAlaIleLysIleAlaSerLeuValValPheAlaLeuLeu 508
QY 1439 GGGTTCACCTTCTACGCGCTGCGATCCTGCGCTACACACTGGCTCTCTCTCTCTCTCT 1498
Db 509 GlyPheProLeuSerIleThrTyrSerValProPheSerIleThrAla----- 524
QY 1499 GAGAGCAGGTCTTCTGCGCCAAATACCGAGGGGACACTGGAGT 1543
Db 525 -----GluLeuThrAlaAspThrGlyGly 532

```

## RESULT 9

```

041972
ID 041972 PRELIMINARY; PRT: 706 AA.
AC 041972;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Hypothetical 66.4 kDa protein.
GN GAMMAHV. M10B.
OS murid herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=33708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WUMS;
RX MEDLINE=97366649; PubMed=9223479;
RA Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,
RA Dal Canto A.J., Speck S.H.;

```

```

RT "Complete sequence and genomic analysis of murine gammaherpesvirus
RL J. Virol. 71:5894-5904(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WUMS;
RA Latreille P., Wamsley P., Waterston R.H.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U97553; AAB66421.1; -
KW Hypothetical protein.
SQ SEQUENCE 706 AA; 66413 MW; 9AD7A4AFDEB8748B CRC64;

```

```

Alignment Scores:
Pred. No.: 1.84e-15 Length: 706
Score: 365.50 Matches: 272
Percent Similarity: 29.39% Conservative: 41
Best Local Similarity: 25.54% Mismatches: 278
Query Match: 5.87% Indels: 474
DB: 12 Gaps: 61

US-09-759-143-110 (1-3410) x 041972 (1-706)
QY 2939 GGAAGTGGGGGAACACAGGCTGGCCAAAGAGAGGGGTGGTTAGGAAGCGGTTGAGA 2880
Db 3 GlyThrGlyGly-----ThrGlyProArgSerGluGly---AlaGlyCysProAlaArg 19
QY 2879 CCTGAAGCCCGCCACCTCTACCTTCTTCAACACCCCTAACCTTGGGTAAACAGCATTTGGAA 2820
Db 20 GlyAlaGlyProProAla----- 25
QY 2819 TTATCATTTGGGATGAGTAGAATTTCCAAGTCTCTGGTTAGGCATTTTGGGGGGCCAGA 2760
Db 26 -----TrpGlyAlaGly 29
QY 2759 CCCCAGGAGAGAGATTTCTGCAATGATCAGCCCAATGACCACTATCTCAGGGGACCT 2700
Db 30 ProProArgArg----- 33
QY 2699 GATTGTTGGGATGCCCGCCACCTACCCAAATATTAGACACACAGAAAAGCTAGCAA 2640
Db 33 ----- 33
QY 2639 TGGATTCCTTCTTACTTTGTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2580
Db 33 ----- 33
QY 2579 GATGGCAACAGAGGACCAACAGGCCACATCTGTATATAAAGGTAAGAGGGGGGTGGATCA 2520
Db 34 AspGlyGlyAsnArg-----GlyAspGly 41
QY 2519 GCATAAAGACAGTGTGTGGGCTGAGGGGACCT-----GGTCTTGTGTGTGGTCCCTC 2466
Db 42 AlaProGluArgGlySerGlyLeuProGlyProGlyCysGlySerProGlyLeuGlySer 61
QY 2465 AGGACTCTTCCCTACAAATAAGTCATATGTTCCAAATCCCATGCGAGAGGTGTTTCATCCT 2406
Db 62 ArgSerProProAlaGly-----TrpGlyGluProGlyGly----- 73
QY 2405 AGAACTCCCATGCAAGAGCTACATTAAACGAAGCTTCAAGGGGCTTAGAGATGG 2346
Db 74 ArgGlyProGlyAlaArg-----GluArgAlaAlaArg-ProGlyValArgValPr 90
QY 2345 GAAACACAGTGTAGTGTATTATTACGCTCCCAAAACCCCTTCTCTAGGTGTGTCTCACT 2286
Db 90 oArgProGlyGluProVal-----ProProGlyGlyMetGlyGlyTh 104
QY 2285 AGGAGGCTAGCTCTTAACCTGAGCTGGGTAAATCCACCTGCAGAGTCCCGCCAGTCCAG 2226
Db 104 rGlyGly-----ThrGlyProArgSerG 112
QY 2225 TGCATGGAGCCCTTCTGGCTCCCTGTATAGTCAGACTGAAACCCCTTGGAGGCT 2166
Db 112 uGly-----AlaGlyCysProAla-ArgGlyAlaGlyProProAlaTrpGlyAlaG 129

```



QY 2639 TGAATCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATCCCTGTGTCTCTGT 2580  
Db 61 ----- 61  
QY 2579 GATGGCAACAGAGACCAACAGCCACATCTCTGATAAAGGTAAAGAGGGGTGGATCA 2520  
Db 62 AspGlyGlyAsnArg -----GlyAspGly 69  
QY 2519 GCAAAAGACAGTGTCTGTGGCTGAGGGACCT-----GGTCTCTGTGTGTGTTCCCTC 2466  
Db 70 AlaProGluArgGlySerGlyLeuProGlyCysGlySerProGlyLeuGlySer 89  
QY 2465 AGGACTCTCCCTTACAAATAAGTCATATGTTCAATCCCATGAGGAGGTGTTTCATCCT 2406  
Db 90 ArgSerProProAlaGly-----TrpGlyGluProGlyGly----- 101  
QY 2405 AGAAACTCCATGCAAGAGCTACATTAAACGAAGCTCAGGTTAAAGGGCTTAGAGATGG 2346  
Db 102 ArgGlyProGlyAlaArg-----GluArgAlaAlaArg-ProGlyValArgValPr 118  
QY 2345 GAAACAGGTGACTGAGTTATTTCAGCTCCCAAAACCTTCTCTAGGTGTGTCTCAACT 2286  
Db 118 oArgProGlyGluProVal-----ProProGlyGlyMetGlyGlyTh 132  
QY 2285 AGGAGGTAGCTGTAAACCTGAGCCTGGGTAATCCACTCCAGAGTCCCGGCATWCCAG 2226  
Db 132 rGlyGly-----ThrGlyProArgSerG1 140  
QY 2225 TGCATGGAGCCTTCTGCTCCTCTGTATAGTCCAGACTCAAAACCTTGGAGGCT 2166  
Db 140 uGly-----AlaGlyCysProAla-ArgGlyAlaGlyProProAlaTrpGlyAlaG 157  
QY 2165 CCAGTCAGCAGCCTAGAGACTGGGAGAGAGAGAGGGACGCT----- 2121  
Db 157 lyProProArgArg---AspGlyGlyAsnArgGlyAspGly-AlaProGluArgGlySer 175  
QY 2120 -----CCAGCCCCCAGTGTGACGTACGACCTCAGCAGCAGAGGTGGCAGCAGAGA 2067  
Db 176 GlyLeuProGlyProGlyCysGlySerProGlyLeuGlySer----- 189  
QY 2066 GCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCGAGCCCATG- 2013  
Db 190 -----ArgSerProProAlaGlyTrpGlyGlu 198  
QY 2012 -----GGCTAACAGAGAGGGGAGCT-----GGACCCAGTGTAGGC 1977  
Db 199 ProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgValPro 218  
QY 1976 AGGCCCTCCACCCCAATGTGCTGGNAGTTTCTACGCTGAGTATTGGCCAACTCGCTCT 1917  
Db 219 ArgProGlyGluProVal----- 224  
QY 1916 TGTCAAATACTACCTGTGTAGCAAAAGTAAATGGCGACCAAGCCAGAGCTCGGCAGACA 1857  
Db 225 -----ProProGlyGlyMetGlyGlyThr 232  
QY 1856 -----CCATATAGCAGTGTACAGACTGGC 1833  
Db 233 GlyGlyThrGlyProArgSerGlyGlyAlaGlyCysProAlaArg----- 247  
QY 1832 TGAGCTGGACATGAGGCCCAATAACAGGATGGGGCCACTGGGACAGCAGGAGGAC 1773  
Db 248 -----GlyAlaGlyProProAlaTrpGlyAlaGlyProProArgArgAspGly-GlyAs 265  
QY 1772 TATCCAGGATGGCAGGTCCAGGAGATGCC-----CCGGCCCGGAACCAACCC 1725  
Db 265 nArgGlyAspGlyAlaProGluArgGlySerGlyLeuProGlyProGlyCysGlySerPr 285  
QY 1724 TGGCTCGGTGGCTCACCACCAACAGTACGAGACATCACAGCAGAGGCCCGC 1665  
Db 285 oGlyLeuGlySerArgSerProProAlaGlyTrpGlyGluProGlyGlyArgGlyProG1 305

QY 1664 AGAGCCG-----GGTGGAGGTGGGAGCAGCCACTGCCTCCAGCAC 1623  
Db 305 yAlaArgGluArgAlaAlaArgProGlyValArg-ValProArgProGlyGluProValP 325  
QY 1622 CCACGTGTCTATTAGGAGGAGCTCCAGGCTTGGGCTTGGCAGGAGCTGGTCATCA 1563  
Db 325 roProGlyGlyMetGly-----GlyThrGlyGlyThrGlyPro---Arg-SerGlyGlyAla 342  
QY 1562 GGCTGTCTCTACTGTGTAGCACCTCCAGTGTCCCTCGTATTGGCAGCAACACTGCT 1503  
Db 343 GlyCysProAla-----ArgGlyAlaGlyProProAlaTrpGlyAlaGlyProProArg 360  
QY 1502 TCTCCCGGTGGT---AGAGGAGGCGCAGTGTAGGGCAGGATCTCAGGCGTGAAGG 1446  
Db 361 ArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySer---GlyLeu----- 377  
QY 1445 TGAACCCGTGAGGGCGGTGAAGCTGCACACAGCCACACTGTGGGACA---GCGATG 1389  
Db 378 -----ProGlyProGlyCysGlySerProGlyLeu 387  
QY 1388 TGGCACCGCAGCCACAGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTCCGA 1329  
Db 388 GlySerArgSerPro----- 392  
QY 1328 ATCGCTGCACCGCTCCATGACCAGAGAGAGCAGGAGATGG---GCGACTGCA 1272  
Db 393 -----ProAlaGly-----TrpGlyGluProGly 400  
QY 1271 GGAACAGCCCGAGCTGCCCATCCGAACGCTTTCATCATAGTGTCTCCGGCTCGGTGC 1212  
Db 401 GlyArgGlyProGlyAla---ArgGluArgAla----- 410  
QY 1211 CCGGCTCAGCTTGGCAGCAGCTTGTGTACA---GCCCTCGCCACAAATCCGTGTAAA 1155  
Db 411 -----AlaArgProGlyValArgValProArgProGlyGluPro----- 423  
QY 1154 ACAGCGTGAAGTCACTAGTGCCA-----TCCAGCTGCACAGCTCAGCCA 1110  
Db 424 -----ValProGlyGlyMetGlyGlyThrGlyGlyThrGlyPro 437  
QY 1109 CGAAGACCGCGCAGGCTCGGGGCA-----TGGCGGAGCAGCAGCTGGTCAGCGGG 1056  
Db 438 ArgSerGluGlyAlaGlyCysProAlaArgGlyAlaGlyProProAlaTrpGlyAlaGly 457  
QY 1055 -----GAAGCAGGGCCCGCAGGTTCGGAAAGCCAAAGC----- 1023  
Db 458 ProProArgArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySerGlyLeu 477  
QY 1022 ---GGCCCGCATGACAGCAGTGGCGGCAAGAGGAGGGCGGCGCAGACCC----- 975  
Db 478 ProGlyProGlyCysGlySerProGlyLeuGlySerArgSerProProAlaGlyTrpGly 497  
QY 974 -----CTTCTGCTGGCTCGGTGGGGCCCGCGCTG 945  
Db 498 GluProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgVal 517  
QY 944 CCT-----CCTACGCCACAGCAGTGTGCTCTACGACAGGTGAGGAAGATGA 897  
Db 518 ProArgProGlyGluProValProProGly----- 527  
QY 896 GGCTGAGCAGCCCAAGAGGCACTCTCCT-----GGTGGCCCGAGTAGG 852  
Db 528 GlyMetGlyGlyThrGlyGlyThrGlyProArgSerGlyGlyAlaGlyCysProAlaArg 547  
QY 851 GGGCAGGCGACTGTGTCTCCAGTCAATGAGGAGGAGCA----- 813  
Db 548 GlyAlaGly-----ProProAlaTrpGlyAlaGlyProProArgArgAspGlyGly 564  
QY 812 ---GGTAGCCAGGCGAGCCCCCAAGACTGATCAAGGATAGAGAGTAGGCGCTGCG 756  
Db 565 AsnArgGlyAspGlyAlaProGluArgGlySer-----Gly 576  
QY 755 GACAGTGGTCCGGT-----CCCGGAAGAGGTTCAGAGAGCAGGCGCTCCAGTGGAG 705



Db 219 AlaCysGlyTyrLeuValGlyAlaMetAspTrpGlyHisSerValLeuGlyArgLeuLeu 238  
Qy 860 GGACACAGGAGGAGTCCCTCTTT-----GGCTGTCTCACCTC 898  
Db 239 GlySerGlyGlnValIleTyrPhePheSerAlaLeuThrTrpGlyValPheLeuIle 258  
Qy 899 ATCTTCCTCACCTCGCTA----- 916  
Db 259 ValHisLeuPheSerIleProGluLysProLeuAlaLysValProSerGluSerSerAla 278  
Qy 917 ---GCACCCACAGTGTGGTGGTGGAGGCA-----CGCTGGGC----- 955  
Db 279 SerSerAlaLeuArgLeuGlyProHisSerAsnGlyTyrGlyAlaLeuGlyLysGlu 298  
Qy 956 CCCACCCAGCAGCAGAGGCTGTCGGCCCTCTCTGTCGCCCTGCTGCTCCATGC 1015  
Db 299 ProValSerProValIleProThrSerSerProGluIleArgProArgSerTyrSerAla 318  
Qy 1015 -----CGGCCCGCTTGGCTTCCGGACCTGGCGCCCTG 1015  
Db 319 LeuGlyGluArgProArgSerPheSerAlaLeuGlyGluAlaAsnSerValThrSerSer 338  
Qy 1016 -----CGGCCCGCTTGGCTTCCGGACCTGGCGCCCTG 1051  
Db 339 AlaLysGlnProIleLysGluAspGlnLysLysMetThrPheArgSer----- 354  
Qy 1052 CTTCCCGCTGCACACAGCTGTGTCGCGCATGCGCCGACCTGCGCGCTCTTCGTG 1111  
Db 355 -----LeuMetLysAlaIlePheAsnMetProAsnHisTyrArgPheLeuCysIle 371  
Qy 1112 GCTGAGCTGTGCAGCTGGATGACCTATGACCTTCCAGCTGTTTACACGGATTCGTG 1171  
Db 372 SerHisLeuLeuGlyTyrAlaAlaPheLeuCysAsnMetLeuPhePheThrAspPheMet 391  
Qy 1172 GGCAGGGCTGTACACAGCGCTGCCAGAGCTGAGCGGGCCAGCGCCCGGAGACAC 1231  
Db 392 GlyGlnIleValTyrArgGlyAsnProTyrAlaGluHisAsnSerThrAlaTyrIleThr 411  
Qy 1232 TATGATGAAGCGCTTCGATGGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1291  
Db 412 TyrGluArgGlyValGluValGlyCysTrpGlyLeuCysIleAsnAlaValSerSerAla 431  
Qy 1292 GTCTTCTCTGTGTGACGCGCTGTGTGACGAGTTCGCGACTCGAGAGCTCTATTG 1351  
Db 432 LeuTyrSerTyrValGlnArgPheLeuLeuProTyrIleGlyLeuLysGlyLeuTyrPhe 451  
Qy 1352 GCCAGTGTGCAGCTTCCCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC 1411  
Db 452 MetGlyTyrPheValPheGlyMetGlyThrSerLeuIleGlyLeuPheProGluValIle 471  
Qy 1412 GTGTGTACAGCTTCAGCGCGCTTCAGCGCTTCACCTTCTCAGCGCTGCAGATCCTGCC 1471  
Db 472 AlaThrLeuIleLeuCysSerValPheGlyValMetSerSerThrLeuTyrThrIlePro 491  
Qy 1472 TACACACTGCTCTCTTACACCGGGAGAGAGAGTGTCTGCTGCCCAATACCGAGG 1531  
Db 492 PheAsnLeuIleAlaGluTyrGlnArgGluGluGluGlnVal---LysLeuGluGly 510  
Qy 1532 GACACTGGAGTGTACAGTGGAGACAGCTGTATGACAGCTTCCTGCGCAGCGCTTAAG 1591  
Db 510 ----- 510  
Qy 1592 CTTGGAGCTCCCTTCCCTAATGGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1651  
Db 511 -----GlyAsnGluSerProArgGlyThrGlyMet----- 520  
Qy 1652 CCACCGCGCTCTGCGGGGCTCTGCTGTGATGTCTCGTACGTGTGTGTGTGTGTGTGT 1711  
Db 521 -----AspCysAlaAlaLeuThrCysMetValGlnLeu----- 531  
Qy 1712 CCACCGAGGCGAGGTGTTCGGCGCGGGGATCTGCTGCGACCTCGCCATCTCGAT 1771  
Db 532 -----AlaGlnIleIleValGlyAlaGly----- 539

Qy 1772 AGTGCCTCTCTGCTGTCCAGGTGGCCCATCCCTGTTTATGGGTCTCATTTGTCGCTC 1831  
Db 540 -----LeuGlyAlaLeuValAsnAla 546  
Qy 1832 AGCCAGCTGTCTCAGTCCCTATATGTGTCTGCCGAGGCTGGGTCTGTCGCCCATTTAC 1891  
Db 547 AlaGlySerValIleValValLeuSerAlaSerSerIleSerLeuIleGlyCysIle 566  
Qy 1892 TTT 1894  
Db 567 Phe 567  
RESULT 7  
ID 041971 PRELIMINARY; PRT; 774 AA.  
AC 041971;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Hypothetical 73.2 kDa protein.  
GN GAMAHV\_M10A.  
OS murid herpesvirus 4.  
OC Viruses; GSDNA viruses, no RNA stage; Herpesviridae;  
OC Gamaherpesvirinae.  
OX NCBI\_TaxID=33708;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WUMS;  
RX MEDLINE=97366649; PubMed=9223479;  
RA Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,  
RA Dal Canto A.J., Speck S.H.;  
RT "Complete sequence and genomic analysis of murine gammaherpesvirus  
68".  
RL J. Virol. 71:5894-5904(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WUMS;  
RA Latreille P., Wamsley P., Waterston R.H.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U97553; AAB66420.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 774 AA; 73164 MW; D1B37F264C92C42C CRC64;  
Alignment Scores:  
Pred. No.: 1 5e-16 Length: 774  
Score: 382.50 Matches: 285  
Percent Similarity: 29.90% Conservative: 43  
Best Local Similarity: 25.98% Mismatches: 291  
Query Match: 6.14% Indels: 478  
DB: 12 Gaps: 67  
US-09-759-143-110 (1-3410) x 041971 (1-774)  
Qy 2939 GGAAGTGGGGGGAACACAGCTGGCCCAAGAGAGGGGTGCTTAGGAAGCCGTGAGA 2880  
Db 31 GlyThrGlyGly-----ThrGlyProArgSerGluGly---AlaGlyCysProAlaArg 47  
Qy 2879 CCTGAAGCCCCACCTCTACCTTCCTTCACACCCTAACCTTGGGTGAACAGCATTTGNA 2820  
Db 48 GlyAlaGlyProAla----- 53  
Qy 2819 TTATCATTTGGGATGAGTAGAGATTTCCAAAGTCTGGTGGTGGTGGTGGTGGTGGT 2760  
Db 54 -----TrpGlyAlaGly 57  
Qy 2759 CCCAGGAGAGAGAGATTTCTGGCAATGATCAGCCCAATGACAGCATCTCTCAGGGGACCT 2700  
Db 58 ProProArgArg----- 61  
Qy 2699 GATTGTTGGGATCCCCCACCCTTACCCAAATATTAGACACACACACAGGCTAGCAA 2640  
Db 61 ----- 61



RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman K.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003552; AAF50310.1;  
DR FlyBase; FBgn0035968; CG4484.  
SQ SEQUENCE 599 AA; 66057 MW; C5381D334CFBF2EB CRC64;

Alignment Scores:  
Pred. No.: 2,13e-25 Length: 599  
Score: 519.50 Matches: 150  
Percent Similarity: 39.43% Conservative: 100  
Best Local Similarity: 23.66% Mismatches: 201  
Query Match: 8.09% Indels: 183  
DB: 5 Gaps: 14

US-09-759-143-110 (1-3410) x 09SV1 (1-599)

QY 317 CGGACGGGAACCCAGCTCTGCTGGTGGCAACCTGTAACCTTTGGCTGGAGGTGT 376  
DB 46 ArgLysThrArgPheGluMetPheArgLeuSerAlaIleAlaMetAlaIleGluPheAla 65  
QY 377 TTGGCGGAGCATCATGATGTCGCGCTCTCTGCTGGAAGTGGGGTAGAGAGAAG 436  
DB 66 TyrAlaAlaGluThrSerPheValSerProIleLeuLeuGlnIleGlyValAspHisLys 85  
QY 437 TTGATGACCATGTGTGGCATGTGTCAGTCTGGCTGGCTGGCTGGCTGGCTGGCT 496  
DB 86 HisMetSerMetThrTrpGlyLeuSerProIleGlyPhePheMetSerProLeuLeu 105  
QY 497 GGCTCAGCAGTACCATGCGCTGGAGCAGCTATGGCGCGCGCGCGCTCATCTGGGCA 556  
DB 106 GlySerIleSerAspArgCysLysLeuArgTrpGlyArgArgProIleIleSerIle 125  
QY 557 CTGCTCTGGCATCTGCTGAGCGCTCTTCTCATCCCA----- 595  
DB 126 LeuSerPheGlyIleMetCysGlyLeuIleLeuValProTyrClyLysAspLeuGlyLeu 145  
QY 596 -----AGG 598  
DB 146 LeuLeuGlyAspAlaGlyTyrThrTyrAlaGluSerAlaLeuAsnPheThrSerSer 165  
QY 599 GCGGGTGGCTAGCAGGGCTGCTGTC-----CCGATCCCGAGGCC 640  
DB 166 GlyClySerValAlaAlaLeuValSerGlyGluAlaThrThrGlyProSerAlaSerAsp 185  
QY 641 CTGGAGCTGGCA-----CTGCTCATCTGGGGTGGGGTGGCTGGCTGGCTGGCT 694  
DB 186 TyrLysPheAlaValIleLeuThrIleLeuGlyMetValLeuLeuAspPheAspAlaAsp 205  
QY 695 GTGTGCTTACTCCATGGAGGCCCTCTCTGTAAGCTCTTCGGGACCGGACCACTGT 754  
DB 206 ThrCysGlnThrProAlaArgThrTyrLeuLeuAspMetCys---ValProGluGluGln 224  
QY 755 CGCAGGCCCTACTCTCTATGCTTTCATGATCATGATCAGTCTTGGGGCTGGCTACCTC 814

DB 225 ProLysAlaMetThrMetPheAlaLeuPheAlaGlyPheGlyGlyThrIleGlyTyrAla 244  
QY 815 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874  
DB 245 IleGlyGlyValAspTrpGluThrThrIleGlySerPheMetGlyGlyAsnIlePro 264  
QY 875 TGCCTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928  
DB 265 ThrValPheThrLeuValThrIlePheAlaValCysTyrLeuIleThrValThrThr 284  
QY 929 -----CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958  
DB 285 PheArgGluIleProLeuProLeuIleGluGlnAspGluLeuLeuArgProLeuSerGlu 304  
QY 958 ----- 958  
DB 305 GlnAlaIleLysLysGluLeuLysLysAsnAsnThrIleTyrTyrIleGlnGluThr 324  
QY 959 -----ACGAGCCAGCAGAGGG 976  
DB 325 ThrGlnLeuGluLeuGlnMetAlaSerAspProLysArgLeuGluAlaLeuGlnGly 344  
QY 977 CTGCTGCGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000  
DB 345 SerTyrGlnAsnGlyTyrSerProAlaValGluLysGlnGlyLysSerGlnAspLeuGlu 364  
QY 1001 -----CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048  
DB 365 ThrGlnSerAspTyrAspAlaProValSerLeuLysAlaTyrLeuLysSerIlePheIle 384  
QY 1049 CTGCTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1108  
DB 385 -----MetProTyrSerMetArgMetLeuAla 393  
QY 1109 GTGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1168  
DB 394 LeuThrAsnLeuPheCysTrpMetGlyHisValThrTyrCysLeuTyrPheThrAspPhe 413  
QY 1169 GTGGCGAGGGCTGTACACAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1228  
DB 414 ValGlyGluAlaValPheHisGlyAspProThrAlaAlaProAsnSerGluAlaLeu 433  
QY 1229 CACTATGATGAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1288  
DB 434 AsnTyrGluAlaGlyValArgPheGlyCysTrpGlyMetAlaIleTyrAlaPheSerCys 453  
QY 1289 CTGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348  
DB 454 SerIleTyrSerLeuSerValThrLysLeuMetLysTrpPheGlyThrLysAlaValTyr 473  
QY 1349 TTGGCAGTGTGGCAGCTTTCCTGCTG-----GCTGCGGTGCCACATGCTGCTGCTGCT 1399  
DB 474 IleSerGlyMetIleTyrGlyIleGlyMetLeuValLeuGlyLeuTrpProThrLys 493  
QY 1400 CACAGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1459  
DB 494 TrpGlyValLeuValPheSerThrSerAla-----GlyIleLeuTyrGlyThrIle 510  
QY 1460 CAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1519  
DB 511 PheThrValProPheIleLeuValAlaArgTyrHis---AlaLysAsnCysPheSerIle 529  
QY 1520 AAATACCGAGGGGACACTGGAGTGTGTAGCAGTGTAGGACAGCTGTGATGATGATGATGATGAT 1579  
DB 530 Lys----- 530  
QY 1580 CCAGGCGCTAAGCCTGGAGCT 1639  
DB 531 -----AsnGly----- 532  
QY 1640 CTGCTCCCACTCCACCGCGCTCTGCGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1699  
DB 532 ----- 532

## Alignment Scores:

Pred. No.: 4,73e-155 Length: 501  
 Score: 2531.00 Matches: 489  
 Percent Similarity: 98.20% Conservative: 2  
 Best Local Similarity: 97.80% Mismatches: 9  
 Query Match: 39.44% Indels: 0  
 DB: 6 Gaps: 0

US-09-759-143-110 (1-3410) x Q95KC5 (1-501)

QY 440 ATGACCATGGCTGGCGCATGGTCCACTGCTGGCGCTGGTCTGTCTCCGCTCCTAGGC 499  
 Db 1 MetThrMetValLeuGlyIleGlyProValLeuGlyLeuValSerValProLeuLeuGly 20  
 QY 500 TCAGCAGTACCTGCTGGCGGAGCGGTATGGCGCGCGCGCGCTTCACTTGGGCACTG 559  
 Db 21 SerAlaSerAspHisTrpArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg 40  
 QY 560 TCCTTGGGCATCTGCTGAGCTCTTCTCATCCCAAGGGCGGCTGGCTAGCAGGCTG 619  
 Db 41 SerLeuGlyIleLeuLeuSerLeuPheLeuIleProArgAlaGlyTrpLeuAlaGlyLeu 60  
 QY 620 CTGTGCCCGGATCCAGGCCCTGGAGCTGGCTGCTCATCTCTGGCGCTGGGCTGCTG 679  
 Db 61 LeuGlyProAspProArgProLeuGlyLeuAlaLeuLeuLeuLeuGlyValGlyLeuLeu 80  
 QY 680 GACTTCTGTGGCCAGGTGCTTTCACCTCCACTGGAGGCGCTGCTCTGACCTCTTCCGG 739  
 Db 81 AspPheCysGlyGlnValCysPheThrProLeuGluAlaLeuLeuSerAspLeuPheArg 100  
 QY 740 GACCGGACCACTGCTGGCAGGCTTCTGCTGCTATGCTTCATGCTTCATGATCATGCTG 799  
 Db 101 AspProAspHisCysArgGlnAlaTy:SerValTrpThrPheMetIleSerLeuGlyGly 120  
 QY 800 TGCCTGGCTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 859  
 Db 121 CysLeuGlyTrpLeuLeuProAlaIleAspTrpAspThrSerAlaLeuAlaProTrpLeu 140  
 QY 860 GGCACCCAGGAGGTGCTCTTTGGCTGTCTCACCCTCATCTTCTCCACTGCTGCTAGCA 919  
 Db 141 GlyThrGlnGluGlyCysLeuPheGlyLeuLeuThrLeuLeuPheLeuThrCysValAla 160  
 QY 920 GCCACACTGCTGGTGGCTGAGGAGGAGCGCTGGCGCCCGCCAGCAGCAGAGGCTG 979  
 Db 161 AlaThrLeuLeuValAlaGluGluAlaLeuGlyProAlaGluProAlaGluGlyLeu 180  
 QY 980 TCGGCGCCCTCTGCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1039  
 Db 181 SerAlaProSerLeuProSerHisCysCysProCysTrpAlaArgLeuAlaPheArgAsn 200  
 QY 1040 CTGGGCGCCCTGCTTCCCGGCTGCACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1099  
 Db 201 LeuGlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArg 220  
 QY 1100 CGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159  
 Db 221 ArgLeuPheValAlaGluLeuCysSerTrpMetAlaLeuMetThrPheThrLeuPheTrp 240  
 QY 1160 ACGATTCTGCTGGGCGAGGCTGCTTACACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1219  
 Db 241 ThrAspPheValGlyGlyGlyLeuTy:GlnGlyValProArgAlaGluLeuGlyThrGlu 260  
 QY 1220 GCCCGGAGACACTATGATGAAGGCTTCCGATGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1279  
 Db 261 AlaArgArgHisTrpAspGluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCys 280  
 QY 1280 GCCATCTCCCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1339  
 Db 281 AlaIleSerLeuValPheSerLeuValValAspArgLeuValGlnArgPheGlyThrArg 300  
 QY 1340 GCAGTCTATTGGCAGTGGGAGCTTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1399

Db 301 AlaValTrpLeuAlaSerValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSer 320  
 QY 1400 CACAGTGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1459  
 Db 321 HisSerValAlaValValThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeu 340  
 QY 1460 CAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1519  
 Db 341 GlnIleLeuProTrpThrLeuAlaSerLeuTy:HisArgGluArgGlnValPheLeuPro 360  
 QY 1520 AATACCGAGGACACTGGAGTGTCTAGCAGTGAGGACAGCTGATGACAGCTGCTGCTGCTGCTGCTG 1579  
 Db 361 LysTrpArgGlyAspAlaGlyThrSerSerGluAspSerLeuMetThrSerPheLeu 380  
 QY 1580 CAGAGCCCTAAGCTGAGCTCCCTTCCCTAATGGACAGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1639  
 Db 381 ProGlyProLysProGlyAlaProPheProAsnGlyHisValGlyAlaGlySerGly 400  
 QY 1640 CTGCTCCACCTCCACCCCGCTGCTGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1699  
 Db 401 LeuLeuProProProAlaLeuCysGlyAlaSerAlaCysAspValSerValArgVal 420  
 QY 1700 GTGGTGGTGGAGCCCGAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1759  
 Db 421 ValValGlyGluProThrGluAlaArgValProGlyArgGlyIleCysLeuAspLeu 440  
 QY 1760 GCATCCTGATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1819  
 Db 441 AlaIleLeuAspSerAlaPheLeuLeuSerGlnValAlaProSerLeuPheMetGlySer 460  
 QY 1820 ATTGCTCCAGCTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1879  
 Db 461 IleValGlnLeuSerGlnSerValThrAlaTy:MetValSerAlaAlaGlyLeuGlyLeu 480  
 QY 1880 GTGCGCATTTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1939  
 Db 481 ValAlaIleTrpPheAlaThrGlnValPheAspLysSerAspLeuAlaLysTrpSer 500

RESULT 4  
 Q9VSV1 ID Q9VSV1 PRELIMINARY; PRT; 599 AA.  
 AC Q9VSV1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE CG4484 protein.  
 GN CG4484.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballwey R.M., Basu A., Baxendale J.P., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

## Alignment Scores:

Pred. No.: 2,98e-172 Length: 553  
 Score: 2798.00 Matches: 542  
 Percent Similarity: 98.55% Conservative: 2  
 Best Local Similarity: 98.19% Mismatches: 8  
 Query Match: 43.60% Indels: 0  
 DB: 6 Gaps: 0

US-09-759-143-110 (1-3410) x Q95KI5 (1-553)

QY	284	ATGGTCCAGAGCTGTGGGTGAGCGCCCTGCTGGCGGACCGAAGCCAGCTTGTGCTG	343
Db	1	MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu	20
QY	344	GTCACCTCTAACCTTTCCGCTGGAGGTGTGTTGGCGCAGGACATCACCTATGTGCGG	403
Db	21	IleAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro	40
QY	404	CCTCTGCTCTGGAAGTGGGGGTAGAGGAGAAGTTTCATGACCATGGTGTGGGCAATTGGT	463
Db	41	ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly	60
QY	464	CCAGTCTGGGCTGTGTGCTGCTCCGCTCCTAGGCTAGCCAGTACGACCTGGCGTGA	523
Db	61	ProValLeuGlyLeuValSerValProLeuLeuGlySerAlaSerAspHisTrpArgGly	80
QY	524	CGCTATGGCGCGCGCGCCCTTCATCTGGGCACCTGTCTTGGGCATCTGCTGAGCCTC	583
Db	81	ArgTyrGlyArgArgProPheIleIrpAlaLeuSerLeuGlyIleLeuLeuSerLeu	100
QY	584	TTTCTCATCCCAAGGCGCGCTGGCTAGCAGGGCTGCTGTGCCGGATCCCAAGGCCCTG	643
Db	101	PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu	120
QY	644	GAGCTGGCACTGCTCATCTCGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	703
Db	121	GluLeuAlaLeuLeuLeuGlyValGlyLeuLeuAlaSerPheCysGlyGlnValCysPhe	140
QY	704	ACTCCACTGGAGCGCTGCTGCTGACCTCTTCCGGACCCCGGACCTGTGCGCAGGCC	763
Db	141	ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla	160
QY	764	TACTGTCTATGCTTATGATCATCTTGGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCTG	823
Db	161	TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla	180
QY	824	ATTGACTGGACACAGTCCCTGGCCCTACCTGGGACCCAGGAGAGTGCCTCTTT	883
Db	181	IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe	200
QY	884	GGCTGCTCACCCTCATCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	943
Db	201	GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu	220
QY	944	GCAGCCTGGGCCCCACGAGCAGCAGAGGGCTGTGGGCCCTCTCTTGTGGCCCGAC	1003
Db	221	AlaAlaLeuGlyProAlaGluProAlaGluGlyLeuSerAlaProSerLeuProSerHis	240
QY	1004	TGCTGTCCATGCGGGCGCGCTTGGCTTTCGGAAACCTGGGCGCGCTGCTCCCGGCTG	1063
Db	241	CysCysProCysTrpAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu	260
QY	1064	CACCACTGTGCTGGCGCTGCGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1123
Db	261	HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys	280
QY	1124	AGCTGATGCGACTCATGACCTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1183
Db	281	SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu	300
QY	1184	TACCAAGGCGTGCACAGACTGAGCGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCG	1243
Db	301	TyrGlnGlyValProArgAlaGluLeuGlyThrGluAlaArgArgHisTyrAspGluGly	320

QY	1244	GTTCCGATGGCAGCTGGGCTGTCTCTGCACTGGCGCATCTCCCTGGTCTTCTCTCTG	1303
Db	321	ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu	340
QY	1304	GTCATGGACCGCTGGTGCACGAGTTCGGCACTCGCAGCTATTTTGGCAGTGTGGCA	1363
Db	341	ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla	360
QY	1364	GCTTTTCCCTGTGGTGGCGGTGCCACATGCTGTCCACAGTGTGGCGGTGGTGCACAGT	1423
Db	361	AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValThrAla	380
QY	1424	TCAGCGCGCTCACCGGTTTCCCTTCTACGCTTCGAGATCTCCCTACACACTGGCC	1483
Db	381	SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla	400
QY	1484	TCCCTCTACCAACCGGAGAGCAGGTGTCTTCCCAATAACGAGGGGACACTGGAGGT	1543
Db	401	SerLeuTyrHisArgGluArgGlnValPheLeuProLysTyrArgGlyAspAlaGlyGly	420
QY	1544	GCTACGACTGAGGACAGCCTGATGACCACTTCTTCCGAGCGCTTCCAGCTTGGAGTCCC	1603
Db	421	ThrSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro	440
QY	1604	TTCCCTTAATGACACAGCTGGGTGCTGGAGCAGTGGCTGCTCCACCTCCACCCGCGCTC	1663
Db	441	PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu	460
QY	1664	TGCGGGGCTCTGCTGTGATGTCTCCGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1723
Db	461	CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla	480
QY	1724	AGGTGTGTTCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1783
Db	481	ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu	500
QY	1784	CTGTCCAGGTGGCGCCCTCCCTCTTATGGGCTCCATTTGCTCCAGCTCAGCAGTCTGT	1843
Db	501	LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal	520
QY	1844	ACTGCTATATGTGTCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1903
Db	521	ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln	540
QY	1904	GTAGTATTGTACAGAGCAGCTTGGCCAAATACTCA	1939
Db	541	ValValPheAspLysSerAspLeuAlaLysTyrSer	552

RESULT 3  
 Q95KC5 PRELIMINARY; PRT; 501 AA.  
 ID Q95KC5 PRELIMINARY; PRT; 501 AA.  
 AC Q95KC5  
 DT 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE Hypothetical 53.4 kDa protein.  
 OS Macaca fascicularis (Cebus eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=MEDULLA OBLONGATA;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries."  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB062977; BAB0745.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 501 AA; 53447 MW; 8C554BBD04EE0470 CRC64;

Pred. No.: 2,59e-176 Length: 553  
 Score: 2861.00 Matches: 553  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 44.58% Indels: 0  
 DB: 4 Gaps: 0

US-09-759-143-110 (1-3410) x Q96JT2 (1-553)

QY 284 ATGGTCCAGAGCTGTGGGTGAGCGCTGCTGGCGGACCGAAGCCAGCTTCTGCTG 343  
 Db 1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeu 20  
 QY 344 GTCAACCTGCTAACCTTTGGCTGGAGGTGTGTTGGCGGAGGATCACCATTATGTGGC 403  
 Db 21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaGlyIleThrTyrValPro 40  
 QY 404 CCTCTGCTCTGGAAGTGGGTAGAGGAGAGTTCATGACCATGGCTGGGATTTGGT 463  
 Db 41 ProLeuLeuGluValGlyValGluLysPheMetThrMetValLeuGlyIleGly 60  
 QY 464 CCAGTGTGGGCTGTGTGCTGCTCCCGCTCTAGGCTCAGCAGTACACTGGCGTGA 523  
 Db 61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80  
 QY 524 CGCTATGGCGCGCGCGCTTCACTGCGGACTGTCTTGGGCACTCCTTGGGCACTCCTGCTGAGCGTC 583  
 Db 81 ArgTyrGlyArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100  
 QY 584 TTTCTCATCCCAAGGCGCGGTGGCTAGCAGGCTGTGCTCCGGATCCCGGATCCCGAGCCCTG 643  
 Db 101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuLysProAspProArgProLeu 120  
 QY 644 GAGCTGGCACTGCTCATCTGGGCTGGGCTGTGGACTTCTGTGGCACTGCTGCTGCTTC 703  
 Db 121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140  
 QY 704 ACTCACTGGAGCGCTCTCTGACCTCTTCCGGGACCGGACCACTGTGCGCAGGCC 763  
 Db 141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgProAspPheHisCysArgGlnAla 160  
 QY 764 TACTCTGTCTATGCTCATCTGCTGCTGGGCTGGCTGGCTGCTGCTGCTGCTGCTGCTG 823  
 Db 161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180  
 QY 824 ATTGACTGGGACACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883  
 Db 181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluCysLeuPhe 200  
 QY 884 GGCCTGCTCACCTCATCTTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943  
 Db 201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220  
 QY 944 GGAGCGTGGGCGCCACCGAGCCAGCAGAGGCTGTGCGGCCCTCTTGTGCGCCAC 1003  
 Db 221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240  
 QY 1004 TGCTGTCCATCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1063  
 Db 241 CysCysProCysArgAlaArgAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260  
 QY 1064 CACAGCTGTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1123  
 Db 261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280  
 QY 1124 ACCTGGATGGCACTATGACCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1183  
 Db 281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300  
 QY 1184 TACCAGGCGTGGCAGAGCTGAGCGGCGGACCGGCGGACACTATGATGAGGC 1243  
 Db 301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

QY 1244 GTTCGGATGGCGAGCCCTGGGCTGTTCTGTCAGTGGCCCATCTCCCTGGTCTTCTCTG 1303  
 Db 321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340  
 QY 1304 GTCATGGACCGGCTGGTGCAGCGATTCCGACACTCGAGCAGTCTATTGGCCAGTGGCA 1363  
 Db 341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360  
 QY 1364 GCTTTCCTGCTGGCTGGCGTGCACATGCTGTCCACAGTGTGGCCGTGGTACAGCT 1423  
 Db 361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380  
 QY 1424 TCAGCCCGCTCAGCGGTTCCACTTCTCAGCCCTGCAGATCTCTCCCTACACACTGGCC 1483  
 Db 381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400  
 QY 1484 TCCTCTTACCACCGGAGAGAGAGTGTCTCTCCCAAAATACCGAGGACACTGGAGGT 1543  
 Db 401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420  
 QY 1544 GCTAGCAGTGCAGACAGCTGATGACAGCTTCTTCCAGGCGCTTAAAGCTGGAGTCCC 1603  
 Db 421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440  
 QY 1604 TTCCTTAATGACACAGCTGGTGTGGAGGAGTGGCTGTCTCCACCTCCACCCGCGCTC 1663  
 Db 441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProAlaLeu 460  
 QY 1664 TGGGGGCTCTGCTGTGTGTGTCTCGTACGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1723  
 Db 461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480  
 QY 1724 AGGTGTGTCGCGCGCGGCGCATCTGCTGGACCTCGCCATCTCGATAGTGTGCTTCTG 1783  
 Db 481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500  
 QY 1784 CTGTCCAGTGTGGCGCCATCCCTGTTTATGGCTCCATTTGTCAGCTCAGCCAGTGTGTC 1843  
 Db 501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520  
 QY 1844 ACTGCTATATGCTGTGCTGCGGAGCGCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTG 1903  
 Db 521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540  
 QY 1904 GTAGTATTGACAGAGCGACTTGGCCAAATACTACGCG 1942  
 Db 541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

# RESULT 2

Q95KI5 ID Q95KI5 PRELIMINARY; PRT; 553 AA.  
 AC Q95KI5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical 59.4 kDa protein.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 ON NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TEMPORAL LOBE RIGHT;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain CDNA  
 RT libraries".  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB060851; BAB46871.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 553 AA; 59392 MW; 0718F3A91FB3BF1E CRC64;

1	2861	44.6	553	4	Q96JT2	Q96jt2 homo sapien
---	------	------	-----	---	--------	--------------------





QY 1417 GACAGCTTCAGCGCCCTCACC GGTTACACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476  
|||||  
Db 215 GACAGCTTCAGCGCCCTCACC GGTTACACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 156  
|||||  
QY 1477 ACTGGCCTCCCTCTACACCGGGAGAGCAGGTGTCCT 1515  
|||||  
Db 155 ACTGGCCTCCCTATACCACCGGGAGAGCAGGTAGTCAT 117  
|||||

Search completed: June 9, 2003, 18:34:21  
Job time : 8556 secs

```

RESULT 43
ARL12290
LOCUS       ARL12290
DEFINITION  Sequence 11 from patent US 6130043.
ACCESSION  ARL12290
VERSION     ARL12290.1 GI:14092190
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 272)
AUTHORS    Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N.,
            Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D.,
            Roberts-Rapp, L., Russell, J.C., Stroupe, S.D. and Yu, H.
TITLE      Reagents and methods useful for detecting diseases of the prostate
JOURNAL    Patent: US 6130043-A 11 10-OCT-2000;
FEATURES   Location/Qualifiers
            source
            1..272
            /organism="unknown"
BASE COUNT 58 a 96 c 49 g 68 t 1 others
ORIGIN
Query Match 7.9%; Score 270; DB 6; Length 272;
Best Local Similarity 99.6%; Pred. No. 2.8e-36;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2795 AAATTCTACTATCCCAATGATAATCCAAATGCTGTGTACCAAGGTTAGGTTGTA 2854
Db 1 AAATTCTACTATCCCAATGATAATCCAAATGCTGTGTACCAAGGTTAGGTTGTA 60

Qy 2855 GGAAGGTAGAGGTGGGGTTCAGGTCTCAAGGGTTCCTCAACCCCTCTTCTCTG 2914
Db 61 GGAAGGTAGAGGTGGGGTTCAGGTCTCAAGGGTTCCTCAACCCCTCTTCTCTG 120

Qy 2915 GCCAGCTGGTTCGCCCACTTCCACTCCCTCTACTCTCTAGACTGGGCTGATGA 2974
Db 121 GCCAGCTGGTTCGCCCACTTCCACTCCCTCTACTCTCTAGACTGGGCTGATGA 180

Qy 2975 AGGCATGCCCCAAATTTCCCTTACCCCACTTCCCTTACCCCACTTCCCACT 3034
Db 181 AGGCATGCCCCAAATTTCCCTTACCCCACTTCCCTTACCCCACTTCCCACT 240

Qy 3035 GCTCCCAACCCCTGTTGGAGCTACTGTCAGG 3065
Db 241 GCTCCCAACCCCTGTTGGAGCTACTGTCAGG 271

RESULT 44
ARL12288
LOCUS       ARL12288
DEFINITION  Sequence 9 from patent US 6130043.
ACCESSION  ARL12288
VERSION     ARL12288.1 GI:14092188
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 265)
AUTHORS    Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N.,
            Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D.,
            Roberts-Rapp, L., Russell, J.C., Stroupe, S.D. and Yu, H.
TITLE      Reagents and methods useful for detecting diseases of the prostate
JOURNAL    Patent: US 6130043-A 9 10-OCT-2000;
FEATURES   Location/Qualifiers
            source
            1..265
            /organism="unknown"
BASE COUNT 64 a 96 c 61 g 76 t
ORIGIN
Query Match 7.8%; Score 265; DB 6; Length 265;
Best Local Similarity 100.0%; Pred. No. 2e-35;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2453 GGGGAAGAGTCTCTGAGGGCAACACACAGAACAGGTCCTCCCTCAGCCACAGCACTGTC 2512
Db 1 GGGGAAGAGTCTCTGAGGGCAACACACAGAACAGGTCCTCCCTCAGCCACAGCACTGTC 60

Qy 2513 TTTTGTGTATGACACCCCTCTTACCTTTTATCAGATGTCGCTTGGTCTTCTCTG 2572
Db 61 TTTTGTGTATGACACCCCTCTTACCTTTTATCAGATGTCGCTTGGTCTTCTCTG 120

Qy 2573 TGCATCAGACAGACACAGGCAATTTAATATTTAACTTATTTTAAACAAGTAGAAG 2632
Db 121 TGCATCAGACAGACAGGCAATTTAATATTTAACTTATTTTAAACAAGTAGAAG 180

Qy 2633 GAATCCATGCTAGCTTTCTGTGCTGCTCTAATATTTGGGTAGGTTGGGATCCCC 2692
Db 181 GAATCCATGCTAGCTTTCTGTGCTGCTCTAATATTTGGGTAGGTTGGGATCCCC 240

Qy 2693 AACATCAGTCCCTCAGATAGCT 2717
Db 241 AACATCAGTCCCTCAGATAGCT 265

RESULT 45
HSP54C06/c
LOCUS       HSP54C06
DEFINITION  H.sapiens flow-sorted chromosome 1 HindIII fragment, SCIP54C06,
            sequence tagged site.
ACCESSION  AL033941
VERSION     AL033941.1 GI:3893713
KEYWORDS   STS; Anonymous marker; single read.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 406)
AUTHORS    Gregory, S., Kettleborough, R., Langford, C., Ross, M.T. and Hunt, S.E.
TITLE      Direct Submission
JOURNAL    Submitted (18-NOV-1998) The Sanger Centre, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
            humquery@sanger.ac.uk
COMMENT    Vector: pBSIISK+
            Marker: stSG3426FS (Primer A: TGTAGGCGAGGATCTGCAG; Primer B:
            CTCGAGGAGCTCTATTGGGC; amplicon size: 141 bp) was mapped to
            chromosome 1 using Radiation Hybrid
            panel Genebridge 4 (GB4).
FEATURES   Location/Qualifiers
            source
            1..406
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clone="SCIP54C06"
            /sex="Female"
            /tissue_type="EBV lymphoblastoid cell line"
            /clone_lib="SCIP"
            /dev_stage="adult"
BASE COUNT 97 a 128 c 115 g 66 t
ORIGIN
Query Match 7.7%; Score 261.4; DB 11; Length 406;
Best Local Similarity 96.1%; Pred. No. 8.4e-35;
Matches 268; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1237 TGAAGCGTTCCGATGGGCGAGCTGGGCTGTTCTCTGAGTGCAGCATCTCCCTGGTCTT 1296
Db 395 TGAAGCGTTCCGATGGGCGAGCTGGGCTGTTCTCTGAGTGCAGCATCTCCCTGGTCTT 336

Qy 1297 CTCTCTGTCATGAGCCGCTGGTGCAGCATCTCGGCACTTCGAGCAGTCTATTTCGCCAG 1356
Db 335 CTATGTGGTTCATGGACCGCTGGTGCAGCATCTCGGCACTTCGAGCAGTCTATTTCGCCAG 276

Qy 1357 TGTGGCAGCTTTCCCTGTGGCTGCCGTCGCCACATGCCCTGTCCACAGTGTGGCGTGT 1416
Db 275 TGTGGCAGCATTCCTCTGTGGCTGCCGTCGCCACATGCCCTGTCCACAGTGTGGCGTGT 216

```

## RESULT 40

AR112287  
LOCUS AR112287 342 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 8 from patent US 6130043.  
ACCESSION AR112287  
VERSION AR112287.1 GI:14092187  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 342)  
AUTHORS Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N., Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D., Roberts-Rapp, L., Russell, J.C., Stroupe, S.D., and Yu, H.  
TITLE Reagents and methods useful for detecting diseases of the prostate  
JOURNAL Patent: US 6130043-A 8 10-OCT-2000;  
FEATURES Location/Qualifiers  
source 1..342  
BASE COUNT 84 a 88 c 81 g 88 t 1 others  
ORIGIN

Query Match 9.4%; Score 319.8; DB 6; Length 342;  
Best Local Similarity 98.5%; Pred. No. 8.6e-45;  
Matches 333; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 2207 GCCAGAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAG 2266  
DB 1 GCCAGAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAG 60  
QY 2267 GGTAAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAAGGTTTTTTGGAGCTGAATA 2326  
DB 61 GGTAAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAAGGTTTTTTGGAGCTGAATA 120  
QY 2327 AACTCAGTCACCTGTTTCCCATCTCTAAGCCCTTAACCTGCCAGCTTCGTTTAATGTAG 2386  
DB 121 AACTCAGTCACCTGTTTCCCATCTCTAAGCCCTTAACCTGCCAGCTTCGTTTAATGTAG 180  
QY 2387 CTCCTGATGGAGTTTCTAGGATGAACACTCTCCCTCCATGGGATTTGAACATATG--ACT 2444  
DB 181 CTCCTGATGGAGTTTCTAGGATGAACACTCTCCCTCCATGGGATTTGAACATATGAAAGT 240  
QY 2445 TATTTGTAGGGAAGAGTCCTGAGGGGCAACACACAGCAAGCAGTCCCTCAGGCCACA 2504  
DB 241 TATTTGTAGGGAAGAGTCCTGAGGGGCAACACACAGCAAGCAGTCCCTCAGGCCACA 300  
QY 2505 GCACGTCTCTTTTGTGATCCACCCCTCTTACCTTT 2542  
DB 301 GCACGTCTCTTTTGTGATCCACCCCTCTTACCTTT 338

## RESULT 41

AR112291  
LOCUS AR112291 294 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 12 from patent US 6130043.  
ACCESSION AR112291  
VERSION AR112291.1 GI:14092191  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 294)  
AUTHORS Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N., Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D., Roberts-Rapp, L., Russell, J.C., Stroupe, S.D., and Yu, H.  
TITLE Reagents and methods useful for detecting diseases of the prostate  
JOURNAL Patent: US 6130043-A 12 10-OCT-2000;  
FEATURES Location/Qualifiers  
source 1..294  
BASE COUNT 70 a 87 c 60 g 74 t 3 others  
ORIGIN

Query Match 8.5%; Score 291; DB 6; Length 294;  
Best Local Similarity 99.0%; Pred. No. 7.1e-40;  
Matches 291; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2974 AAGGCACGTGCCCAAAATTTCCCTACCCCAACTTTCCCTACCCCAACTTTCCCAACC 3033  
DB 1 AAGGCACGTGCCCAAAATTTCCCTACCCCAACTTTCCCTACCCCAACTTTCCCAACC 60  
QY 3034 AGCTCCACACACCTGTTTGGAGCTACTGCAGGACAGCAAGCAAAAGTCGGTTTCCCAA 3093  
DB 61 AGCTCCACACACCTGTTTGGAGCTACTGCAGGACAGCAAGCAAAAGTCGGTTTCCCAA 120  
QY 3094 GCCTTTGTCATCTCAGCCCCCAGAGATATCTGTGTGGGGAATCTCACACAGAACT 3153  
DB 121 GCCTTTGTCATCTCAGCCCCCAGAGATATCTGTGTGGGGAATCTCACACAGAACT 180  
QY 3154 CAGGAGCACCCCTGCTGAGCTAAGGAGGTCTTATCTCAGGGGGGTTTAAGTGCC 3213  
DB 181 CAGGAGCACCCCTGCTGAGCTAAGGAGGTCTTATCTCAGGGGGGTTTAAGTGCC 240  
QY 3214 GTTTGCAATAATGTCGTCTTATTTATTAGCGGGGTGAATATTTTATCTGTA 3267  
DB 241 GTTTGCAATAATGTCGTCTTATTTATTAGCGGGGTGAATATTTTATCTGTA 294

## RESULT 42

AR112289  
LOCUS AR112289 288 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 10 from patent US 6130043.  
ACCESSION AR112289  
VERSION AR112289.1 GI:14092189  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 288)  
AUTHORS Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N., Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D., Roberts-Rapp, L., Russell, J.C., Stroupe, S.D., and Yu, H.  
TITLE Reagents and methods useful for detecting diseases of the prostate  
JOURNAL Patent: US 6130043-A 10 10-OCT-2000;  
FEATURES Location/Qualifiers  
source 1..288  
BASE COUNT 70 a 66 c 61 g 90 t 1 others  
ORIGIN

Query Match 8.4%; Score 287; DB 6; Length 288;  
Best Local Similarity 99.7%; Pred. No. 3.4e-39;  
Matches 287; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2532 CTCCTACCTTTTATCAGGATGTGCGCTGTGGTCTCTCTGTCCTATCAGAGACACAG 2591  
DB 1 CTCCTACCTTTTATCAGGATGTGCGCTGTGGTCTCTCTGTCCTATCAGAGACACAG 60  
QY 2592 GCATTTAAATATTTAACTTATTTTAAACAAAGTAGAAGGGAATCATTTGCTAGCTTT 2651  
DB 61 GCATTTAAATATTTAACTTATTTTAAACAAAGTAGAAGGGAATCATTTGCTAGCTTT 120  
QY 2652 CTGTGTTGGTCTTAATATTTTGGGTGGGATGCCCAACAAATCAGGTCCTCTGAG 2711  
DB 121 CTGTGTTGGTCTTAATATTTTGGGTGGGATGCCCAACAAATCAGGTCCTCTGAG 180  
QY 2712 ATAGCTGGTCATTTGGCTGATCATTCGCCAAATCTTCTCTCTGGGGTCTGGCCCCCA 2771  
DB 181 ATAGCTGGTCATTTGGCTGATCATTCGCCAAATCTTCTCTCTGGGGTCTGGCCCCCA 240  
QY 2772 AAATGGCTTAACCCAGGACCTTGGAAATTTCTACTCATCCCCCAATGATAA 2819  
DB 241 AAATGGCTTAACCCAGGACCTTGGAAATTTCTACTCATCCCCCAATGATAA 288









RESULT 37  
AX267037/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

QY 2072 GCTGCCACCCCTGTGCTGAGGTGCGTGTGAGTGCACAGCTGGGGGCTGGGGCGTCCCTCT 2131  
|||||  
Db 615 GCTGCCACCCCTGTGCTGTGAGGTGCGTGTGAGTGCACAGCTGGGGGCTGGGGCGTCCCTTT 556  
QY 2132 CTTCTCTCCAGCTCTAGGGCTGCGTGTGAGTGCACAGCTGGGGGCTGGGGCGTCCCTCT 2191  
|||||  
Db 555 CTTCTCTCCAGCTCTAGGGCTGCGTGTGAGTGCACAGCTGGGGGCTGGGGCGTCCCTCT 496  
QY 2192 GACTTATACAGGGGCGGAGGCTCCATGCACTGGAATGCGGGGACTCTGCAAGTGG 2251  
|||||  
Db 495 GACTTATACAGGGGCGGAGGCTCCATGCACTGGAATGCGGGGACTCTGCAAGTGG 436  
QY 2252 ATTACCCAGGCTCAGGGTTACAGCTAGCTCCCTAGTTGAGACACACTAGAGAAGGGTT 2311  
|||||  
Db 435 ATTACCCAGGCTCAGGGTTACAGCTAGCTCCCTAGTTGAGACACACTAGAGAAGGGTT 376  
QY 2312 TTTGGGAGCTGAATAAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2371  
|||||  
Db 375 TTTGGGAGCTGAATAAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316  
QY 2372 CTTCTCTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2431  
|||||  
Db 315 CTTCTCTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256  
QY 2432 TGAACATATG--ACTTATTTAGGGGAGAGTCTGAGGAGGAGTCTGAGGAGGAGTCTGAGG 2489  
|||||  
Db 255 TGAACATATGAAAGTATTTAGGGGAGAGTCTGAGGAGGAGTCTGAGGAGGAGTCTGAGG 196  
QY 2490 TCCCTCTAGCCACAGCACTGCTTTTGTGCTGATCCACCCCTGCTTACCTTTATCAGG 2549  
|||||  
Db 195 TCCCTCTAGCCACAGCACTGCTTTTGTGCTGATCCACCCCTGCTTACCTTTATCAGG 136  
QY 2550 ATGTGGCTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2609  
|||||  
Db 135 ATGTGGCTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 76  
QY 2610 TATTTATTTAAAGTGAAGGAGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2669  
|||||  
Db 75 TATTTATTTAAAGTGAAGGAGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 16  
QY 2670 TTTGGGTAGGGTGGG 2684  
|||||  
Db 15 TTTGGGTAGGGTGGG 1

RESULT 35  
AX140521/c  
LOCUS AX140521  
DEFINITION Sequence 11 from Patent WO0134802.  
ACCESSION AX140521  
VERSION AX140521.1 GI:14280639  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,  
Reed, S.G., Kalos, M.D., Retter, M.W., Stolk, J.A., Day, C.H.,  
Skelky, Y.A. and Wang, A.  
TITILE Compositions and methods for the therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0134802-A 11 17-MAY-2001;  
CORIXA CORPORATION (US)  
FEATURES source  
1. .772  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 221 a 202 c 170 g 162 t 17 others  
ORIGIN  
Query Match 17.7%; Score 603; DB 6; Length 772;

Best Local Similarity 91.8%; Pred. No. 3.6e-93;  
Matches 675; Conservative 0; Mismatches 49; Indels 11; Gaps 4;  
QY 1961 TTGGGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2015  
|||||  
Db 735 TTGGGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676  
QY 2016 TTGGGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2071  
|||||  
Db 675 TTGGGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616  
QY 2072 GCTGCCACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2131  
|||||  
Db 615 GCTGCCACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556  
QY 2132 CTTCTCTCCAGCTCTAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2191  
|||||  
Db 555 CTTCTCTCCAGCTCTAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496  
QY 2192 GACTTATACAGGGGCGGAGGCTCCATGCACTGGAATGCGGGGACTCTGCAAGTGG 2251  
|||||  
Db 495 GACTTATACAGGGGCGGAGGCTCCATGCACTGGAATGCGGGGACTCTGCAAGTGG 436  
QY 2252 ATTACCCAGGCTCAGGGTTACAGCTAGCTCCCTAGTTGAGACACACTAGAGAAGGGTT 2311  
|||||  
Db 435 ATTACCCAGGCTCAGGGTTACAGCTAGCTCCCTAGTTGAGACACACTAGAGAAGGGTT 376  
QY 2312 TTTGGGAGCTGAATAAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2371  
|||||  
Db 375 TTTGGGAGCTGAATAAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316  
QY 2372 CTTCTCTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2431  
|||||  
Db 315 CTTCTCTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256  
QY 2432 TGAACATATG--ACTTATTTAGGGGAGAGTCTGAGGAGGAGTCTGAGGAGGAGTCTGAGG 2489  
|||||  
Db 255 TGAACATATGAAAGTATTTAGGGGAGAGTCTGAGGAGGAGTCTGAGGAGGAGTCTGAGG 196  
QY 2490 TCCCTCTAGCCACAGCACTGCTTTTGTGCTGATCCACCCCTGCTTACCTTTATCAGG 2549  
|||||  
Db 195 TCCCTCTAGCCACAGCACTGCTTTTGTGCTGATCCACCCCTGCTTACCTTTATCAGG 136  
QY 2550 ATGTGGCTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2609  
|||||  
Db 135 ATGTGGCTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 76  
QY 2610 TATTTATTTAAAGTGAAGGAGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2669  
|||||  
Db 75 TATTTATTTAAAGTGAAGGAGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 16  
QY 2670 TTTGGGTAGGGTGGG 2684  
|||||  
Db 15 TTTGGGTAGGGTGGG 1

RESULT 36  
AX200381/c  
LOCUS AX200381  
DEFINITION Sequence 11 from Patent WO0151633.  
ACCESSION AX200381  
VERSION AX200381.1 GI:15390177  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,  
Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W.,  
Stolk, J.A., Skelky, Y.A., Wang, A. and Meagher, M.J.  
TITILE Compositions and methods for the therapy and diagnosis of prostate  
cancer

\* 73651 81298: contig of 7648 bp in length  
 \* 81299 81398: gap of 100 bp  
 \* 81399 91238: contig of 9840 bp in length  
 \* 91239 91338: gap of 100 bp  
 \* 91339 127752: contig of 36414 bp in length  
 \* 127753 127852: gap of 100 bp  
 \* 127853 177486: contig of 49634 bp in length  
 \* 177487 177586: gap of 100 bp  
 \* 177587 178567: contig of 981 bp in length.

## FEATURES

source

1. .178567  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP23-27H13"  
 /clone\_lib="RPC1-23 Female Mouse BAC"

## misc\_feature

1. .58509  
 /note="assembly\_fragment"

clone\_end:SP6  
 vector\_side:left"

## misc\_feature

58610. .59459  
 /note="assembly\_fragment"

## misc\_feature

59560. .60623  
 /note="assembly\_fragment"

## misc\_feature

60724. .62950  
 /note="assembly\_fragment"

## misc\_feature

63051. .65309  
 /note="assembly\_fragment"

## misc\_feature

65410. .67839  
 /note="assembly\_fragment"

## misc\_feature

67940. .69956  
 /note="assembly\_fragment"

## misc\_feature

70057. .73550  
 /note="assembly\_fragment"

## misc\_feature

73651. .81298  
 /note="assembly\_fragment"

## misc\_feature

81399. .91238  
 /note="assembly\_fragment"

## misc\_feature

91339. .127752  
 /note="assembly\_fragment"

## misc\_feature

127853. .177486  
 /note="assembly\_fragment"

## misc\_feature

177587. .178567  
 /note="assembly\_fragment"

clone\_end:T7  
 vector\_side:right"

BASE COUNT 47757 a 40927 c 41360 g 47318 t 1205 others

## ORIGIN

Query Match 18.3%; Score 623.6; DB 2; Length 178567;

Best Local Similarity 86.8%; Pred. No. 1.7e-96;

Matches 686; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 455 GGCATTGGCTGCTGCTGGCCCTGGTCTGTGCTCCGCTCTAGGCTCAGCCAGTGACCAC 514

DB 171183 GGCATTGGCCAGTGTAGGCCCTGGTCTGTGCTCCACTCTAGGCTCAGCCAGTGACCAC 171124

QY 515 TGGGTGGAGCTATGCGCCCGCCGCGCCCTTCATCTGGSCACTGTCTTGGGCATCCG 574

DB 171123 TGGCGTGGCGCTATGCGCCCGCGAGACCCCTTATCTGGGCTTTGCTCCCTGGGTGCTG 171064

QY 575 CTGAGCTCTTTCTATCCCAAGCGCGGCTGCTAGCAGGGCTGTGTGCGCGGATGCC 634

DB 171063 CTAGGCTCTTTCTATCCCGAGGGCTGGCTGGCTGCTAGCAGGGCTGTGTACCCAGACAC 171004

QY 635 AGCCCTGTGAGCTGACCTGCTATCTGCGGGTGGGCTGCTGACCTCTGTGGCCAG 694

DB 171003 AGCCCTGTGAGTGTGGCCCTGTGCTATCTGGGAGTGGGCTGCTGACCTTTGTGGCCAG 170944

QY 695 GTGTGCTTCACTCACTGGAGGCCCTGCTCTGTGACCTCTTCCGGGACCGGACCACTGT 754

DB 170943 GTGTGCTTCACTCACTGGAGGCCCTTACTCTCGGACCTCTTCCGGGACCGGACCACTGC 170884

QY 755 CGCCAGGCTCTCTCTATGCTTCAATGATGCTTTGGGGGCTGTGCTGGGCTACCTC 814

DB 170883 CGCAAGCCTTCTCTGTCTACGCCCTTCATGATCAGCCTTGGGGCTGCTGGGCTACCTC 170824

QY 815 CTGCTGTCATTTGACTGGGACACACAGTGGCCCTGGCCCTTACTCTGGGACCCAGAGAG 874

DB 170823 TTACCTGGCATTGACTGGGACACACAGCGTTCTGGCCCTTACTCTGGGAAAGAA 170764

QY 875 TGCCTCTTTGGGCTGCTCACCTCATCTTCTCCTCACTGCTGAGTGGTGGTGGTGGTGGTGG 934

DB 170763 TGCCTCTTTGGGCTGCTCACCTCATCTTCTCCTCACTGCTGAGTGGTGGTGGTGGTGG 170704

QY 935 GCTGAGGAGGAGCGCTGGGCCCCACCGAGCCAGCAGAGAGGCTGTGCGCCCTCCCTTGG 994

DB 170703 ACGGAGGAGGAGTACTTGGGCCCCACCGAGCCAGGAGGCTGTGCTCTCTGCCGTG 170644

QY 995 TCGCCCCACTGCTGCTCCATGCGGCGCCGCTTGGCTTCCGGAACCTGGGCGCCCTGCTT 1054

DB 170643 TCGCGCGGATGCTGCCCATGCGCACCTTGGCTTCCGGAATCTGGGTACCCCTGTT 170584

QY 1055 CCGCGGCTGCACAGCTGTGCTGCGCATGCGCCGACCCGACCCCTGCGCGGCTCTTCTGGCT 1114

DB 170583 CCGCGGCTGCACAGCTGTGCTGCGCATGCGCCGACCCCTGCGACCCCTACCGGACTCTTGTGGCT 170524

QY 1115 GAGCTGTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1174

DB 170523 GAGCTGTGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 170464

QY 1175 GAGGGCTGTACAGGGGCTGCGCAGAGCTGAGCGGCGGACCGGACCGGCGGAGACACTAT 1234

DB 170463 GAGGGCTGTACAGGGGCTGACCGAGCGGACCGGACCGGCGGAGACACTAT 170404

QY 1235 GATGAAGCGG 1244

DB 170403 GATGAAGCTG 170394

## RESULT 34

AX106230/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

misc\_feature

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY 875 TGCCCTTTGGCTGCTACCCCTCATCTTCTCCACCTGCGTAGCAGCCACACTGCTGGTG 934  
 Db 161997 TGCCTCTTGGCCCTCCTCACCTCATTTTCTCATCTGTGTGGCAGCCACTGCTGTGTG 162056  
 QY 935 GCTAGAGGAGCGCTGTGGCCCGCCACCGAGCCAGCAGAGGCTGTGGCCCGCTCTCTT 994  
 Db 162057 GCTAGAGGAGCGCTGTGGCCCGCCACCGAGCCAGCAGAGGCTGTGTGGCTCTCTCT 162116  
 QY 995 TCGCCCGCCACTGCTGCTCCATCGCCGCGCTGTGGCTTTCGGAACCTCGCGCGCTGCTT 1054  
 Db 162117 TCAGCGCGGTGCTGCTCATGCGATGCTGCGCTTTCGGAATCTGGGTACCCCTGTTT 162176  
 QY 1055 CCCCGGCTGACAGCTGTGCTGCGCCATGCGCCGACCTGCGCCGCTCTTCGTGGCT 1114  
 Db 162177 CCCCGGCTGACAGCTGTGCTGCGCCATGCGCCGACCTGCGCCGCTCTTCGTGGCT 162236  
 QY 1115 GAGCTGTGACGCTGATGAGGCTCATGACCTTCACGCTGTTTACAGGATTCGTGGCT 1174  
 Db 162237 GAGCTGTGACGCTGATGAGGCTCATGACCTTCACGCTGTTTACAGGATTCGTGGCT 162296  
 QY 1175 GAGGGGTGACAGGCGGTGCCAGAGCTGAGCGGGCCACCGAGGCGCCGAGACACTAT 1234  
 Db 162297 GAGGGGTGACAGGCGGTGCCAGAGCTGAGCGGGCCACCGAGGCGCCGAGACACTAT 162356  
 QY 1235 GATGAGGCG 1244  
 Db 162357 GATGAGGCT 162366

## RESULT 33

## AC107837/c

LOCUS AC107837 179567 bp DNA linear HTG 14-APR-2002  
 DEFINITION Mus musculus clone RP23-272H13, WORKING DRAFT SEQUENCE, 13 ordered pieces.

## ACCESSION

## AC107837

## VERSION

## HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.

## KEYWORDS

## Mus musculus

## SOURCE

## Mus musculus

## ORGANISM

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## 1 (bases 1 to 179567)

## Birren,B., Linton,L., Nusbaum,C. and Lander,E.

## Mus musculus, clone RP23-272H13

## Unpublished

## 2 (bases 1 to 179567)

## Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

## Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,

## Choepeil,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,

## Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,

## Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

## Ginde,S., Gord,S., Goyette,M., Graham,L., Johnson,R., Jones,C.,

## Hagot,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Lamazares,R.,

## Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

## Landers,T., Lechoczy,J., Levine,R., Liu,G., MacLean,C.,

## Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,

## McEwan,V., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,

## Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,

## Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,

## Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,

## Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,

## Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,

## Seaver,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

## Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

## Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,

## Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

## Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

## Direct Submission

## Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome

## Research, 320 Charles Street, Cambridge, MA 02141, USA

## TITLE

## JOURNAL

## COMMENT

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
 Chazaro,B., Choepeil,Y., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagot,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,  
 Landers,T., Lechoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McKernan,K., Meldrum,J., Meneus,L.,  
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
 Seaman,S., Seaver,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information

Center project name: L20597

Center clone name: 272\_H13

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 175041 bases at least Q40

Consensus quality: 176485 bases at least Q30

Consensus quality: 177028 bases at least Q20

Insert size: 182000; agarose-fp

Insert size: 177367; sum-of-contigs

Quality coverage: 7.7 in Q20 bases; agarose-fp

Quality coverage: 7.9 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 13 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have

\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 58509: contig of 58509 bp in length

\* 58510 58609: gap of 100 bp

\* 58610 59459: contig of 850 bp in length

\* 59460 59559: gap of 100 bp

\* 59560 60623: contig of 1064 bp in length

\* 60624 60723: gap of 100 bp

\* 60724 62950: contig of 2227 bp in length

\* 62951 63050: gap of 100 bp

\* 63051 63509: contig of 2259 bp in length

\* 63510 65409: gap of 100 bp

\* 65410 67839: contig of 2430 bp in length

\* 67840 67939: gap of 100 bp

\* 67940 69956: contig of 2017 bp in length

\* 69957 70056: gap of 100 bp

\* 70057 73550: contig of 3494 bp in length

\* 73551 73650: gap of 100 bp

*	1	1063:	contig of 1063 bp in length
*	1064	1163:	gap of unknown length
*	1164	2325:	contig of 1162 bp in length
*	2326	2435:	gap of unknown length
*	2426	3556:	contig of 1131 bp in length
*	3557	3656:	gap of unknown length
*	3657	4752:	contig of 1096 bp in length
*	4753	4832:	gap of unknown length
*	4853	5972:	contig of 1120 bp in length
*	5973	6073:	gap of unknown length
*	6073	7818:	contig of 1746 bp in length
*	7819	7918:	gap of unknown length
*	7919	9266:	contig of 1348 bp in length
*	9267	9366:	gap of unknown length
*	9367	10614:	contig of 1248 bp in length
*	10615	10714:	gap of unknown length
*	10715	11914:	contig of 1200 bp in length
*	11915	12014:	gap of unknown length
*	12015	14075:	contig of 2061 bp in length
*	14076	14175:	gap of unknown length
*	14176	15601:	contig of 1426 bp in length
*	15602	15701:	gap of unknown length
*	15702	17277:	contig of 1576 bp in length
*	17278	17377:	gap of unknown length
*	17378	18657:	contig of 1280 bp in length
*	18658	18757:	gap of unknown length
*	18758	20392:	contig of 1635 bp in length
*	20393	20492:	gap of unknown length
*	20493	21641:	contig of 1149 bp in length
*	21642	21741:	gap of unknown length
*	21742	23412:	contig of 1671 bp in length
*	23413	23512:	gap of unknown length
*	23513	25087:	contig of 1575 bp in length
*	25088	25187:	gap of unknown length
*	25188	26859:	contig of 1672 bp in length
*	26860	26950:	gap of unknown length
*	26960	28698:	contig of 1739 bp in length
*	28699	28790:	gap of unknown length
*	28799	30716:	contig of 1918 bp in length
*	30717	30816:	gap of unknown length
*	30817	32509:	contig of 1693 bp in length
*	32510	32609:	gap of unknown length
*	32610	34220:	contig of 1611 bp in length
*	34221	34320:	gap of unknown length
*	34321	35560:	contig of 1240 bp in length
*	35561	35660:	gap of unknown length
*	35661	37768:	contig of 2108 bp in length
*	37769	37868:	gap of unknown length
*	37869	39562:	contig of 1694 bp in length
*	39563	39662:	gap of unknown length
*	39663	41676:	contig of 2014 bp in length
*	41677	41776:	gap of unknown length
*	41777	43658:	contig of 1882 bp in length
*	43659	43758:	gap of unknown length
*	43759	45624:	contig of 1866 bp in length
*	45625	45724:	gap of unknown length
*	45725	48100:	contig of 2376 bp in length
*	48101	48200:	gap of unknown length
*	48201	50489:	contig of 2289 bp in length
*	50490	50589:	gap of unknown length
*	50590	52329:	contig of 1740 bp in length
*	52330	52429:	gap of unknown length
*	52430	54544:	contig of 2115 bp in length
*	54545	54644:	gap of unknown length
*	54645	56752:	contig of 2108 bp in length
*	56753	56852:	gap of unknown length
*	56853	58639:	contig of 1787 bp in length
*	58640	59739:	gap of unknown length



```
Db 421 CCATCCTGATAGT-CCTTCCCTGCTGTCCANGTGGCCCAFCCTCTGTTATGGGCTCCA 479
QY 1821 TTGTCAGCTCAGCCAGTGTCTACTGCTATATATGTTGTCGCCGAGCCCTGGGTCTGG 1880
Db 480 TTGTCAGCTCAGCCAGTGTCTACTGCTATATATGTTGTCGCCGAGCCCTGGGTCTGG 539
QY 1881 TCGCCATTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTACG 1940
Db 540 TC-CCATTACTTTGCTACACAGGTANTATTGACAAGAAGCANTTGGCCAAATACTACG 598
QY 1941 CGTAGAAACTTCCAGCA--CATTTGGGTGGAGGCGCTGCCCTACTGGGTCCAGCTCCC 1998
Db 599 CGTTAAATAATCCAGCAACTTTGGGGGTGGAGGCGCTGCCCTACTGGGT-CCAACTCCC 657
QY 1999 CGCTCCTGTTAGCCCATGGGCTGCCGGGTGGCCGAGTTTCTGTTGCTGGCCAAAGT 2058
Db 658 CGCTCCTGTTAACCCATGGGCTGCCGGCTGGCCGCGCAATTTCTGTTGCTGGCCAAAT 717
QY 2059 AATGTGGCTCTGCTGCTGCCACCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGC 2117
Db 718 NATGTGGCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANTNGGGGG 777
QY 2118 TGGGGGCTCCC 2128
Db 778 TNGGGNGTCC 788

RESULT 30
AX200380
LOCUS AX200380
DEFINITION Sequence 10 from Patent WO0151633.
ACCESSION AX200380
VERSION AX200380.1 GI:15390176
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
XU, J., DILLON, D.C., MITCHAM, J.L., HARLOCKER, S.L., JIANG, Y.,
REED, S.G., KALOS, M.D., FANGER, G.R., DAY, C.H., RETTER, M.W.,
STOLK, J.A., SKEIKY, Y.A., WANG, A. and MEAGHER, M.J.
TITLE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0151633-A 10 19-JUL-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..789
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 122 a 250 c 221 g 182 t 14 others
ORIGIN

Query Match 19.7%; Score 673.4; DB 6; Length 789;
Best Local Similarity 94.2%; Pred. No. 3.3e-105;
Matches 745; Conservative 0; Mismatches 40; Indels 6; Gaps 5;

QY 1341 CAGTCTATTGCGCAGTGTGGCAGCTTCCCTGTGGCTGCCGTCACATGCTGTGCC 1400
Db 1 CAGTCTATTGCGCAGTGTGGCAGCTTCCCTGTGGCTGCCGTCACATGCTGTGCC 60

QY 1401 ACAGTGTGGCGGTGTGACAGCTTCCAGCCGCCCTCACCAGGTTACCTTTCTCAGCCCTGC 1460
Db 61 ACAGTGTGGCGGTGTGACAGCTTCCAGCCGCCCTCACCAGGTTACCTTTCTCAGCCCTGC 120

QY 1461 ACATCCTGCCCTACACACTGGGCTCCCTCTACCACCGGAGAACGAGGTGTTCTGCCCA 1520
Db 121 AGATCCTGCCCTACACACTGGGCTCCCTCTACCACCGGAGAACGAGGTGTTCTGCCCA 180

QY 1521 AATACCGAGGGGACACTGGAGTGTCTAGCAGTGAGGACAGCTGTATGACCACTGCTCCG 1580
Db 181 AATACCGAGGGGACACTGGAGTGTCTAGCAGTGAGGACAGCTGTATGACCACTGCTCCG 240
```

```
QY 1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGACACAGTGGGTGCTGGAGGCAAGTGCC 1640
Db 241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGACACAGTGGGTGCTGGAGGCAAGTGCC 300
QY 1641 TGTCCCACTCCACCCGCGCTGCGGGGCGCTTGCCTGTGTATGCTCCGTACGTGTGG 1700
Db 301 TGTCCCACTCCACCCGCGCTGCGGGGCGCTTGCCTGTGTATGCTCCGTACGTGTGG 360
QY 1701 TGGTGGGTGAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1760
Db 361 TGGTGGGTGAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 1761 CCATCCTGGATAGTGTCTGCTGCTGCCAGGTGCCAGGTGCCAGGTGCCAGGTGCCAGGT 1820
Db 421 CCATCCTGGATAGT-CTTCTGCTGCCANGTGGCCCATCCCTGTTATGGGCTCCA 479
QY 1821 TTGTCCAGCTCAGCCAGTGTCTACTGCCCTATATGTTGTCGCCGAGGCTGGGTCTGG 1880
Db 480 TTGTCCAGCTCAGCCAGTGTCTACTGCCCTATATGTTGTCGCCGAGGCTGGGTCTGG 539
QY 1881 TCGCCATTACTTTGCTACACAGGTAGTATTTGACAAGAGGAGGAGGAGGAGGAGGAGGAG 1940
Db 540 TC-CCATTACTTTGCTACACAGGTANTATTGACAAGAAGCANTTGGCCAAATACTACG 598
QY 1941 CGTAGAAACTTCCAGCA--CATTTGGGTGGAGGCGCTGCCCTACTGGGTGCCAGCTCCC 1998
Db 599 CGTTAAATAATCCAGCAACTTTGGGGGTGGAGGCGCTGCCCTACTGGGT-CCAACTCCC 657
QY 1999 CGCTCCTGTTAGCCCATGGGCTGCCGGGTGGCCGAGTTTCTGTTGCTGGCCAAAGT 2058
Db 658 CGCTCCTGTTAACCCATGGGCTGCCGGCTGGCCGCGCAATTTCTGTTGCTGGCCAAAT 717
QY 2059 AATGTGGCTCTGCTGCTGCCACCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGC 2117
Db 718 NATGTGGCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANTNGGGGG 777
QY 2118 TGGGGGCTCCC 2128
Db 778 TNGGGNGTCC 788

RESULT 31
AX267036
LOCUS AX267036
DEFINITION Sequence 10 from Patent WO0173032.
ACCESSION AX267036
VERSION AX267036.1 GI:16515821
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
XU, J., DILLON, D.C., MITCHAM, J.L., HARLOCKER, S.L., JIANG, Y.,
KALOS, M.D., FANGER, G.R., RETTER, M.W., STOLK, J.A., DAY, C.H.,
VEDVICK, T.S., CARTER, D., LI, S.X., WANG, A., SKEIKY, Y.A., HEPLER, W.T.
and HENDERSON, R.A.
TITLE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0173032-A 10 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..789
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 122 a 250 c 221 g 182 t 14 others
ORIGIN

Query Match 19.7%; Score 673.4; DB 6; Length 789;
Best Local Similarity 94.2%; Pred. No. 3.3e-105;
Matches 745; Conservative 0; Mismatches 40; Indels 6; Gaps 5;
```





QY	379	GGCGCAGGCATCACCTATGTGGCGCCTCTGCTGCTGGAAGTGGGGGTACAGGAGAAGTT	438
Db	402	GGCGGAATTCATCACCTATGTGGCGCCTCTGCTGCTGGAAGTGGGGGTACAGGAGAAGTT	461
QY	439	CATGACCATGGTGTGGGCATTTGGTCAGTGCCTGGGCTGGTGTCTGCCGTTCCTTAGG	498
Db	462	CATGACCATGGTGTGGGCATTTGGTCAGTGCCTGGGCTGGTGTCTGTGTCCTCCTTAGG	521
QY	499	CTCAGCCATGACCACTGGCGTGAGCGCTATGGCGCGCGCGCCCTTCATCTGGGCACT	558
Db	522	CTCAGCCATGACCACTGGCGTGAGCGCTATGGCGCGCGCGCCCTTCATCTGGGCACT	581
QY	559	GTCTTTGGGCATCCTGCTGAGCCTCTTTCTATCCCAAGGCGCGGTGGCTACGAGGGCT	618
Db	582	GTCTTTGGGCATCCTGCTGAGCCTCTTTCTATCCCAAGGCGCGGTGGCTACGAGGGCT	641
QY	619	GCTGTGCCGGATCCAGAGCCCTCGAGCTGGACCTGCTCATCTGGGCGTGGGGCTGCT	678
Db	642	GCTGTGCCGGATCCAGAGCCCTCGAGCTGGACCTGCTCATCTGGGCGTGGGGCTGCT	701
QY	679	GGACTTCTGTGGCCAGGTGTGCTCACTCCACTGGAGGCCCTGCTCTGACCTCTCCG	738
Db	702	GGACTTCTGTGGCCAGGTGTGCTCACTCCACTGGAGGCCCTGCTCTGACCTCTCCG	761
QY	739	GGACCCGGACACTGTGCCAGGCCCTACTGTCTATGCCCTCATCATCACTCTTGGGG	798
Db	762	GGACCCGGACACTGTGCCAGGCCCTACTGTCTATGCCCTCATCATCACTCTTGGGG	821
QY	799	CTGCTTGGGCTACTCTGCTGCCATTGACTGGGACACCAAGTGCCTTGGCCCCCTACCT	858
Db	822	CTGCTTGGGCTACTCTGCTGCCATTGACTGGGACACCAAGTGCCTTGGCCCCCTACCT	881
QY	859	GGCACCCAGGAGAGTGCTCTTTGGCCTGCTCACCTCATCTCTCATCACTGCGTAGC	918
Db	882	GGCACCCAGGAGAGTGCTCTTTGGCCTGCTCACCTCATCTCTCATCACTGCGTAGC	941
QY	919	AGCCACACTGCTGGCTGTAGAGGCGAGCGCTGGGCGCCACCAGCGACGAGAGGGCT	978
Db	942	AGCCACACTGCTGGCTGTAGAGGCGAGCGCTGGGCGCCACCAGCGACGAGAGGGCT	1001
QY	979	GTGCGGCCCTCCTTGTGCGCCCACTGCTGTCCATGCGCGGCCGCTTGGCTTTCGGAA	1038
Db	1002	GTGCGGCCCTCCTTGTGCGCCCACTGCTGTCCATGCGCGGCCGCTTGGCTTTCGGAA	1061
QY	1039	CTGTGGCGGCTGCTTCCCGGCTGTACCAAGCTGTGCTGCCGATCCCCGACACCTTGGC	1098
Db	1062	CTGTGGCGGCTGCTTCCCGGCTGTACCAAGCTGTGCTGCCGATCCCCGACACCTTGGC	1121
QY	1099	CGGGCTTTCGTGGCTGAGCTGTGCAAGCTGGATGGCACTCATCACTTACGCTGTTTTA	1158
Db	1122	CGGGCTTTCGTGGCTGAGCTGTGCAAGCTGGATGGCACTCATCACTTACGCTGTTTTA	1181
QY	1159	CACGGATTTCTGGGCGAGGG	1179
Db	1182	CACGGATTTCTGGGCGAGTG	1202

RESULT	27
AX267877	
LOCUS	
DEFINITION	AX267877 1203 bp DNA linear PAT 26-OCT-2001
ACCESSION	Sequence 851 from Patent WO0173032.
VERSION	AX267877
KEYWORDS	AX267877.1 GI:16516515
SOURCE	.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y., Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H., Wedick,T.S., Carter,D.R., Lla,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.

Db	2364	CTCAGCCTCTGGGCGATATCTGTCAGAGCCTGGGAATCTCACAC--AAACTCAGGAGTA	2421
QY	3162	CCCCCTGCGTGAAGTAAGGAGAGCTTATCTCTCAGGGGGGTTTAAGTGCCTGTTGCAG	3221
Db	2422	CTCCCGCCAGAGCCAGTCAAGCTCATCAGCTCGA--GGGGGATTTCACTGCTGATTGCAG	2479
QY	3222	TAATGT-----CGTCCTTATTTATTTAGCGGGTGAATATTTTATACITGTAAGTGAGCAA	3275
Db	2480	TANTGTTATCTATGTCCTTATTTATTTAGCAGAGTAATATTTTATACCGTATGTGATAAA	2539
QY	3276	TCAGAGTATAATGTTTATGTGACAAAAATTAAGCGCTTTCCTATATATGTTTAAAAA	3335
Db	2540	TCAGAGTATAATGTTTATGTGATGAAATTAAGC--TTCCGTATATGTTTACTTTGAAAAA	2598
QY	3336	AAAAA	3348
Db	2599	AAAAA	2611

```

RESULT 25
AX343860
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
    1593 bp
    DNA
    linear
    PAT 01-FEB-2002
Sequence 6 from Patent WO0200892.
AX343860.1 GI:18491927
    synthetic construct.
    synthetic construct
    artificial sequences.
1
Cabezon Silva, T.E. and Delisse, A.M.
Triple fusion proteins comprising ubiquitin fused between
thioredoxin and a polypeptide of interest
Patent: WO 0200892-A 6 03-JAN-2002;
SmithKline Beecham Biologics SA (BE)
    Location/Qualifiers
        1..1593
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="Chimaeric (E. coli - human)"
308 a 472 c 461 q 352 t
BASE COUNT

```

Query Match	28.2%;	Score 961.2;	DB 6;	Length 1593;
Best Local Similarity	99.7%;	Pred. No. 2.2e-154;		
Matches 963;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	283	TATGTCACAGAGCTGTGGGTGAGCGCCCTGCTGCGGCACCGGAAAGCCAGCTCTTGCT	342	
Db	603	TATGTCACAGAGCTGTGGGTGAGCGCCCTGCTGCGGCACCGGAAAGCCAGCTCTTGCT	662	
Qy	343	GGTCAACCTGCTAACCTTTGGCCCTGGAGGTGTGTTTGGCCGACAGGCATCACCTATGTGCC	402	
Db	663	GGTCAACCTGCTAACCTTTGGCCCTGGAGGTGTGTTTGGCCGACAGGCATCACCTATGTGCC	722	
Qy	403	GCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTTTCATCACCATGGTCTGGGCATTGG	462	
Db	723	GCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTTTCATCACCATGGTCTGGGCATTGG	782	
Qy	463	TCCAGTGTCTGGSCCTTGCTCTGTGCCGCTCCATAGGCTACGCCAGTGACCATGCGCGTGG	522	
Db	783	TCCAGTGTCTGGSCCTTGCTCTGTGCCGCTCCATAGGCTACGCCAGTGACCATGCGCGTGG	842	
Qy	523	ACGCTATGGCCGCCCGGCCCTTCATCTGGGCACATGCTCTGGGGCATCCTGCTCAGGCT	582	
Db	843	ACGCTATGGCCGCCCGGCCCTTCATCTGGGCACATGCTCTGGGGCATCCTGCTCAGGCT	902	
Qy	583	CTTTTCTCATCCCAAGGGCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCGAGGCCCT	642	
Db	903	CTTTTCTCATCCCAAGGGCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCGAGGCCCT	962	
Qy	643	GGAGCTGGCACATGCTCATCTCTGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTT	702	

963	Db	GGAGCTGGCAC	TGCTCATCTGGGCTGGGGCTGCTGGACTTCTGTGGCCAGGTGCTT	1022
703	QY	CAC	TCCACTGGAGGCCCTGCTCTGACCTCTTCCGGGACCCGAGCACTGTGCGCAGGC	762
1023	Db	CAC	TCCACTGGAGGCCCTGCTCTGACCTCTTCCGGGACCCGAGCACTGTGCGCAGGC	1082
763	QY	CTACTCTGTCTATG	CCCTCATGATACAGTCTTGGGGGCTGCCTGGGTACCTCTGCTGCTGC	822
1083	Db	CTACTCTGTCTATG	CCCTCATGATACAGTCTTGGGGGCTGCCTGGGTACCTCTGCTGCTGC	1142
823	QY	CATTGACTGGGAC	ACCAGTGCCTGGCCCCCTACCTGGGCACCCAGAGAGAGTGCCTCTT	882
1143	Db	CATTGACTGGGAC	ACCAGTGCCTGGCCCCCTACCTGGGCACCCAGAGAGAGTGCCTCTT	1202
883	QY	TGGCCTGCTAC	CCCTCATCTTCTCACCCTGCTAGCAGCCACACTGCTGCTGCTGAGGA	942
1203	Db	TGGCCTGCTAC	CCCTCATCTTCTCACCCTGCTAGCAGCCACACTGCTGCTGCTGAGGA	1262
943	QY	GGCAGGCGCTGGG	CCCCACCAGGACAGAGAAGGGCTGTCGGCCCCCTCTTGTGCGCCCCA	1002
1263	Db	GGCAGGCGCTGGG	CCCCACCAGGACAGAGAAGGGCTGTCGGCCCCCTCTTGTGCGCCCCA	1322
1003	QY	CTGCTGTCCATG	CGGGCCGCTTGGCTTTCGGGAACCTGGGGCCCTGCTTCCCGGCT	1062
1323	Db	CTGCTGTCCATG	CGGGCCGCTTGGCTTTCGGGAACCTGGGGCCCTGCTTCCCGGCT	1382
1063	QY	GCACCACTGTCTG	CGCATGCCGCCACCCCTGCGCCGCTCTTCTCGTGGCTGAGCTGTG	1122
1383	Db	GCACCACTGTCTG	CGCATGCCGCCACCCCTGCGCCGCTCTTCTCGTGGCTGAGCTGTG	1442
1123	QY	CAGCTGATGGCA	CTCATGACCTTTCAGCGCTGTTTTACACGGATTTCTGGCGCAGGGGCT	1182
1443	Db	CAGCTGATGGCA	CTCATGACCTTTCAGCGCTGTTTTACACGGATTTCTGGCGCAGGGGCT	1502
1183	QY	GTACCAGGGCGTG	CCCCAGAGTGAAGCCGGGACCCAGAGCCCGGAGACACTATGATGAAGG	1242
1503	Db	GTACCAGGGCGTG	CCCCAGAGTGAAGCCGGGACCCAGAGCCCGGAGACACTATGATGAAGG	1562
1243	QY	CGTTGC	1248	
1563	Db	CAC	TAG	1568

RESULT	26
AX201078	
LOCUS	AX201078
DEFINITION	Sequence 708 from Patent WO0151633.
ACCESSION	AX201078
VERSION	AX201078.1 GI:15390883
KEYWORDS	human.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1203)
REFERENCE	Xu, J., Dillon, D.C., Mitcham, J. L., Harlocker, S. L., Jiang, Y., Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W., Stolk, J.A., Skeiky, Y.A., Wang, A. and Meagher, M.J. Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL	Patent: WO 0151633-A 708 19-JUL-2001;
FEATURES	CORIXA CORPORATION (US) Location/Qualifiers 1..1203 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	175 a 415 c 368 g 245 t
ORIGIN	
Query Match	23.3%; Score 794.6; DB 6; Length 1203;
Best Local Similarity	99.5%; Pred. No. 6.5e-126;
Matches 797; Conservative	0; Mismatches 4; Indels 0; Gaps 0



```
QY 1220 GCCCGAGACACTATGATGAAGCGTTCGGATGGCGACCTGGGGCTGTTCTTCGACGTGC 1279
Db 1384 CCGCGAGACACTATGATGAAGCGTTCGGATGGCGACCTGGGGCTGTTCTTCGACGTGC 1443
QY 1280 GCCATCTCCCTGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1339
Db 1444 GCCATCTCCCTGGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1503
QY 1340 GCAGTCTATTATTTGGCAGTGTGGCAGCTTTCCTCTGGCTGGCTGGCGTGCACATGCTGTC 1399
Db 1504 GCAGTCTATTATTTGGCAGTGTGGCAGCTTTCCTCTGGCTGGCGTGCACATGCTGTC 1563
QY 1400 CACAGTGTGGCGTGGTGGACAGCTTACAGCCGCTTACAGCGGTTCACCTTCTCAGCCCTG 1459
Db 1564 CACAGTGTGGCGTGGTGGACAGCTTACAGCCGCTTACAGCGGTTCACCTTCTCAGCCCTG 1623
QY 1460 CAGATCTCCCTACACACTGGCTCCCTCTACACCGGGAGAGAGAGAGTGTCTCTGCC 1519
Db 1624 CAGATCTCCCTACACACTGGCTCCCTCTACACCGGGAGAGAGAGAGTGTCTCTGCC 1683
QY 1520 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGGAGGACAGCTGTATGACAGCTTCTG 1579
Db 1684 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGGAGGACAGCTGTATGACAGCTTCTG 1743
QY 1580 CCAGGCTTAAGCTGGAGCTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1639
Db 1744 CCAGGCTTAAGCTGGAGCTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1803
QY 1640 CTGCTCCACCTCCACCGGCTCTGGGGGCTCTGCTGTATGCTCCGTACGCTG 1699
Db 1804 CTGCTCCACCTCCACCGGCTCTGGGGGCTCTGCTGTATGCTCCGTACGCTG 1863
QY 1700 GTGGTGGTGGACCGGACCGAGGAGGTGTTCCGGGCGGGGACATCCCTGTTATGGGCTC 1759
Db 1864 GTGGTGGTGGACCGGACCGAGGAGGTGTTCCGGGCGGGGACATCCCTGTTATGGGCTC 1923
QY 1760 GCCATCTCGATAGTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1819
Db 1924 GCCATCTCGATAGTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1983
QY 1820 ATTGTCCAGCTCAGCAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1879
Db 1984 ATTGTCCAGCTCAGCAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2043
QY 1880 GTGGCATTACTTGTGTACAGAGTAGTATTGACAGAGCGACTTGGGCAATATCTCA 1939
Db 2044 GTGGCATTACTTGTGTACAGAGTAGTATTGACAGAGCGACTTGGGCAATATCTCA 2103
QY 1940 GCG 1942
Db 2104 GCG 2106
```

```
RESULT 24
BC034084 2611 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, clone IMAGE:4457452, mRNA, partial cds.
DEFINITION BC034084
ACCESSION BC034084
VERSION BC034084.1 GI:21707615
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
```

Email: cgabs-re@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK plate: 43 Row: p Column: 14  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.

```
FEATURES             Location/Qualifiers
     source            1..2611
     organism="Mus musculus"
     db_xref="taxon:10090"
     map="FVB/N"
     clone="IMAGE:4457452"
     tissue_type="Mammary tumor. Metallothionien-TGF alpha
     model. 10 month old virgin mouse. Taken by biopsy."
     /clone_lib="NCI_CGAP_Mam1"
     /lab_host="DH10B"
     /note="Vector: pCMV-SPORT6"
     <1..1353
     /codon_start=1
     /product="Unknown (protein for IMAGE:4457452)"
     /protein_id="AAH34084.1"
     /db_xref="GI:21707616"
     /db_xref="LocusID:212980"
     /translation="PRAGLGLYDTPRLPLEALLILGVLLDFCGVCFPLEALL
     SDFRDPDRCROAFSVYAPMISLGGCLVLLPAIDWTSVLAPYLTGTECEFLGLTL
     IELICMAATLFVTEAVLGPPEPAGLLVSAYSRCCPCCHVGLAERNLGTLPRLQOL
     CCMPTGLRLFLVAELCSWALMTFTFTDFVGGELTQGVPRAPGPEARHRSVVT
     RNSGTLFLQCAISLVSLVMDRLVKFQTSVSLASVMTFPAVAAATCLSHSVVVT
     ASAAITGFTFSALQILPTLASLYHREKQVFLPKYRAGDSSGSDSTPLPQPKP
     GALFPNGHVHSGSGGILAPPALCGASACDMSRVVGVPPPEARVVTGRGICLDIAL
     DSAFLLSQVAPSLFMGSIQVLSHVSATYAVMSAAGLGLVAIYFATQVVDKNDLAKYS
     "
```

## CDS

```
BASE COUNT      519 a  728 c  691 g  673 t
ORIGIN
Query Match      36.5%; Score 1246.2; DB 10; Length 2611;
Best Local Similarity 73.1%; Pred. No. 4.3e-203;
Matches 2026; Conservative 0; Mismatches 568; Indels 179; Gaps 26;

QY 593 CCAAGGCGCGCTGGCTAGCAGGCTGTGTGCCCGGATCCAGGCGCTTGGAGTGGCA 652
Db 1 CCGAGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 60

QY 653 CTGCTCATCTGGCGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 712
Db 61 CTGCTCATCTGGCGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 120

QY 713 GAGGCGCTGCTCTCTGACCTCTTCCGGGACCCGCGACCTGTGCGGACGCTACTCTG 772
Db 121 GAGGCGCTTACTCTCCGACCTCTTCCGGGACCCGCGACCTGTGCGGACGCTACTCTG 180

QY 773 TATGCGCTTATGATGATGCTTGTGGGGCTGCTGGGGCTGCTGGGGCTGCTGGGGCTG 832
Db 181 TAGCGCTTATGATGATGCTTGTGGGGCTGCTGGGGCTGCTGGGGCTGCTGGGGCTG 240

QY 833 GACACGAGTCCCTGCCCTTACCTACCTGGGACCCAGGAGGAGTGGCTCTTGGCTGCT 892
Db 241 GACACGAGGCTTCTGGCGGCTTACCTACCTGGGAGTACAGGAGAGTGGCTCTTGGCT 300

QY 893 ACCTCATCTTCTTCTACCTGCTAGCAGGACACACTGCTGGTGGCTGAGGAGGAGCGCTG 952
Db 301 ACCTCATCTTCTTCTTCTGCTAGCAGGACCACTGCTGTTGTGACGAGGAGGAGCTG 360
```





Db 380 GCTGACTCAAGACAGCTGTGGAGTATGTAGTCCCTGGAAACCTTACCTGCGCTGTCC 439  
QY 281 ACTATGCTCAGAGCTGTGGGTAGCGCCTGCTGCGCACCGACCGAAGCCACAGCTCTG 340  
Db 440 ATCATGATCCAGAGCTGTGGCCAGCGCTGTCTACGCGCACCGAAGCTCAGCTCTG 499  
QY 341 CTGGTCAACCTGCTAACCTTTGGCCTGGAGTGTGTTGGCCGAGGACATCACCATTGTG 400  
Db 500 CTGGTCAACCTGCTCACCTTTGGCCTGGAGTGTGCTGCTGGCTGGCTGGCGCATTA 559  
QY 401 CGCCTCTGCTGCTGGAAGTGGGGTAGAGAGAGTTTCAATGACCATGGTGGCTGGGCA 460  
Db 560 CCACCCCTTCTGCTGGAAGTGGGGTGGAGGAGAAATTCATGACCATGGTGTGGGCAT 619  
QY 461 GTTCAGTGTGGCCTGTGCTGTCTCCGCTCTAGCTAGCTCAGCCAGTACCATGGCT 520  
Db 620 GGCCTAGTGTAGGCTGTGTTCTTCCACTCTAGGCTCAGCCAGTGGAGTGGGCT 679  
QY 521 GGACGCTATGGCCGCGCGGCTTCTATCTGGGCACTGTCTTGGGCATCCTGCTGAGC 580  
Db 680 GGGCGCTATGGCCGCGCGGAGACCTTTATCTGGGCTTTGCTCCCTGGGTGCTCTGCT 739  
QY 581 CTCCTTCTCATCCCAAGGCGCGGCTGTAGAGAGGCTGTGTCGCCCGGATCCAGGCC 640  
Db 740 CTCCTTCTCATCCCGAGGCTGGCTGGCTGGCAGGACTGTCTACCCAGACACCGGCC 799  
QY 641 CTGGAGTGGCACTGCTCATCTGCGGCTGGGCTGTGCTGCTGCTGCTGCTGCTGCTG 700  
Db 800 CTGGAGTGGCCCTGCTGATCTTGGAGTGGGGCTGTGCTGCTGCTGCTGCTGCTGCT 859  
QY 701 TTCACTCCACTGGAGGCGCTGCTCTGACCTCTTCCGGGACCGGACCACTGTGCGCAG 760  
Db 860 TTACTCCATTGGAGGCTTACTTCCGACCTCTTCCGGGACCGGACCACTGTGCGCAA 919  
QY 761 GCCTACTGTCTATGCTTCTATGATCAGTCTTGGGGCTGCTGGGCTGCTGCTGCTGCT 820  
Db 920 GCCTTCTCTGTACGCTTCTATGATCAGCTTGGGGCTGCTGGGCTGCTGCTGCTGCT 979  
QY 821 GCATTGCTGGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880  
Db 980 GCATTGCTGGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1039  
QY 881 TTGGGCTGCTCACCCTCATCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 940  
Db 1040 TTGGGCTCCTCACCCTCATCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1099  
QY 941 GAGGAGGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000  
Db 1100 GAGGAGTACTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159  
QY 1001 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060  
Db 1160 CGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219  
QY 1061 CTGCAACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1120  
Db 1220 CTGCAACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1279  
QY 1121 TGCAGCTGATGGCACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180  
Db 1280 TGCAGCTGATGGCACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339  
QY 1181 CTGTACAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1240  
Db 1340 CTGTACAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1399  
QY 1241 GCGTTCGATGGGAGCTGGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1300  
Db 1400 GGCATTCGAATGGGAGCTGGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1459  
QY 1301 CTGGTCACTGGACGCTGGTGGAGGATTCGGCACTTCGGCACTTCGGCACTTCGGCACT 1360  
Db 1473 CTGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2532

Db 1460 CTGCTCATGACAGGCTGTGTACAGAAATTCGGCACAGGCTAGTCTATCTGCGCCAGTGT 1519  
QY 1361 GCAGCTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1420  
Db 1520 ATGACCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1579  
QY 1421 GCTTACAGCCCTTACCCGCTTACCTTTCTGAGCCCTGAGATCTGCTGCTGCTGCTGCTGCT 1480  
Db 1580 GCCTCAGCTGCCCTCACCAGGTTTCACTTCTGCGCTTGCAGATCTGCTTACACGCTC 1639  
QY 1481 GCCTCCTCTTACACCGGAGAGAGGCTTCTTCTGCCCAATACACGAGGAGGACACTGGA 1540  
Db 1640 GCCTCCTCTTACACCGGAGAGGCTTCTTCTGCCCAATACACGAGGAGGAGCTGGA 1699  
QY 1541 GGTGTAGTGTGAGGACAGCTGATGACAGCTTCTGCGAGGCTTAACTGCTGAGCT 1600  
Db 1700 GGTAGCAGGCTGAGGACAGCTGATGACAGCTTCTGCGAGGCTTAACTGCTGAGCT 1759  
QY 1601 CCCTTCCCTTAATGACACAGCTGGGTGTGAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1660  
Db 1760 CTCTTCCCAATGACACAGCTGGGTGTGAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1819  
QY 1661 CTCTGCGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1720  
Db 1820 CTCTGCGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1879  
QY 1721 GCCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1780  
Db 1880 GCCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1939  
QY 1781 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1840  
Db 1940 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1999  
QY 1841 GTCACTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1900  
Db 2000 GTCACTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2059  
QY 1901 CAGTGTATTTGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1960  
Db 2060 CAGTGTATTTGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2117  
QY 1961 TTGGGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2020  
Db 2118 ATCAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161  
QY 2021 CTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2080  
Db 2162 CTGAGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2216  
QY 2081 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2140  
Db 2217 CTAGGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2273  
QY 2141 CCAGTCTTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2200  
Db 2274 CTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2333  
QY 2201 AGGAGGCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2260  
Db 2334 AGGAT-----TGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2373  
QY 2261 GCTCAGGCTTAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2320  
Db 2374 GCTCCAGGCTTAAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2431  
QY 2321 TGAATAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2380  
Db 2432 TGAATAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2472  
QY 2381 ATGTAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2440  
Db 2473 CTGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2532



```
QY 2430 TTTGAACATATGAC--TTATTTGTAGGGAAGAGTCTCTGAGGGGCAACACACACAGAACCA 2487
Db 143150 TTTGAACATATGAAAGTTATTTGTAGGGAAGAGTCTCTGAGGGGCAACACACAGAACCA 143209
QY 2488 GGTCCCTCAGCCACAGCACTGTCTTTTGTCTGATCCACCCCTTACCTTTATCA 2547
Db 143210 GGTCCCTCAGCCACAGCACTGTCTTTTGTCTGATCCACCCCTTACCTTTATCA 143269
QY 2548 GGATGTGGCTGTGTGCTCTCTGTCATCATCAGAGACACAGGCAATTAATATTTAA 2607
Db 143270 GGATGTGGCTGTGTGCTCTCTGTCATCATCAGAGACACAGGCAATTAATATTTAA 143329
QY 2608 CTTATTTATTAACAAAGTAGAAGGAATCCATTGCTAGCTTTCTGTGTGTGTCTAA 2667
Db 143330 CTTATTTATTAACAAAGTAGAAGGAATCCATTGCTAGCTTTCTGTGTGTGTCTAA 143389
QY 2668 TATTTGGTAGGTGGGGGATCCCAACATCAGGTCTCCCTGAGATAGCTGTCTATGGG 2727
Db 143390 TATTTGGTAGGTGGGGGATCCCAACATCAGGTCTCCCTGAGATAGCTGTCTATGGG 143449
QY 2728 CTGATCATTTGCCAATCTTCTCTCTGCTGGGTCTGCCCCCCCCCAAAATGCTTAACCCAGG 2787
Db 143450 CTGATCATTTGCCAATCTTCTCTCTGCTGGGTCTGCCCCCCCCCAAAATGCTTAACCCAGG 143509
QY 2788 ACCTTGAATTTCTATCTATCCCAATATATAATCCAAATGCTGTGTGTGTGTGTGTGT 2847
Db 143510 ACCTTGAATTTCTATCTATCCCAATATATAATCCAAATGCTGTGTGTGTGTGTGTGT 143569
QY 2848 TGTGTAAGGAAGTAGAGGTGGGCTTCAGGTCTCAAGGCTTCCCTTAACCAACCCCTCT 2907
Db 143570 TGTGTAAGGAAGTAGAGGTGGGCTTCAGGTCTCAAGGCTTCCCTTAACCAACCCCTCT 143629
QY 2908 TCTCTTGGCCAGCCTGTGTTCCCTCCCACTTCCACTCCCTCTACTCTCTCTAGGACTGGG 2967
Db 143630 TCTCTTGGCCAGCCTGTGTTCCCTCCCACTTCCACTCCCTCTACTCTCTAGGACTGGG 143689
QY 2968 CTGATGAAGCACTGCCAAATTTCCCTTACCCCACTTTCCCTTACCCCACTTTTC 3027
Db 143690 CTGATGAAGCACTGCCAAATTTCCCTTACCCCACTTTCCCTTACCCCACTTTTC 143749
QY 3028 CCACAGCTCCCAACACCTGTTGGAGTCTACTGCAGGACCAAGCAACAAAGTGGGTT 3087
Db 143750 CCACAGCTCCCAACACCTGTTGGAGTCTACTGCAGGACCAAGCAACAAAGTGGGTT 143809
QY 3088 TCCCAAGCCTTTGCTATCTCAGCCCTCCAGAGTATATCTGTGTTGGGAATCTCACACA 3147
Db 143810 TCCCAAGCCTTTGCTATCTCAGCCCTCCAGAGTATATCTGTGTTGGGAATCTCACACA 143869
QY 3148 GAACTCAGGAGCACCCCTGCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTA 3207
Db 143870 GAACTCAGGAGCACCCCTGCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTA 143929
QY 3208 AGTCCGCTTTGCAATATGTCGCTTATTTATTTAGCGGGTGAATTTTATCTGTAA 3267
Db 143930 AGTCCGCTTTGCAATATGTCGCTTATTTATTTAGCGGGTGAATTTTATCTGTAA 143989
QY 3268 GTGAGCAATCAGAGTATATGTTATGTCGACAAAATTAAGGCTTTCTTATATGTTTA 3326
Db 143990 GTGAGCAATCAGAGTATATGTTATGTCGACAAAATTAAGGCTTTCTTATATGTTTA 144048
```

```
RESULT 22
BC031381
LOCUS BC031381 3354 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, clone MGC:32471 IMAGE:5050610, mRNA, complete cds.
ACCESSION BC031381
VERSION BC031381.1 GI:21594808
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3354)
```

AUTHORS  
TITLE  
JOURNALREMARK  
COMMENT

Strausberg, R.  
Direct Submission  
Submitted (06-JUN-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 45 Row: 1 Column: 2  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.

FEATURES  
source

Location/Qualifiers  
1..3354  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N"  
/clone="MGC:32471 IMAGE:5050610"  
/tissue\_type="Liver, normal. 5 month old male mouse."  
/clone\_lib="NCI\_CGAP\_L19"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
443..2104  
/codon\_start=1  
/product="Unknown (protein for MGC:32471)"  
/protein\_id="AAH31381.1"  
/db\_xref="GI:21594809"  
/db\_xref="LocusID:212980"  
/translation="MIQRLWASRLLRHKAQQLLVNLLTFLGLEVLAAGTYVPPLLL  
EVGVEKEMTWLVGIVGLVPLGSDQWGRYRRRPFIWALSGLVLSLFL  
IPRAGLALGLVPTPLLEALLILGLVLCPCQVCTPLEALLSFLPDPHCRQA  
ESVAFMISLGGCLGILLPAIDWDTSLAPYLQTEBCEFLGLLFLICMAATLFTV  
EAVLGPPEPAGLLVSAVSRRCPCVHGLAFRLNLTFLPRLQCCMPRTLRRLV  
AELCSWALMTFTLFYDFVGEGLYQVPAEPGTEARRHYDEGIRKSLGLFQCAI  
SLVSLVMDRLVQKFTRSVYLSVMTFPVAAAATCLSHSVYVTAALTGFTFSAL  
QILPTLASLVHREKQVFLPKYRGDAGSSGDSQTSFLPGPKPGLFPHGHVSGS  
SGILAPPALCGASACDVMRVVVGPEPPEARVVTGRGICLDLAILDSAFLLSQVAPSL  
FMGSTVQLSHSVTAYMVAAGLVAVIYFATQVVPDNDLAKYSV"

## CDS

BASE COUNT  
ORIGIN

662 a 922 c 934 g 836 t  
Query Match 46.4%; Score 1583.4; DB 10; Length 3354;  
Best Local Similarity 74.0%; Pred. No. 1e-260;  
Matches 2456; Conservative 0; Mismatches 671; Indels 190; Gaps 28;  
QY 43 GCCTCGCCAGATCTGATGATGAGACGTGTCCCTACTGAG--GTGCCCCACAGCAGCA 100  
Db 209 GCGTCCCAAGGATCTGACCGACGAGATGTGTCCCAATCAAGGACCTAGATGGTGA 268  
QY 101 GGTGTTGAGCATGGCTCAGAGCTGGACCGGCCACCAAGCGCTGGCAAAATGGCGCC 160  
Db 269 COTGTTTAGCGTGGGACGAGATGCTGAATTTGACCTAAAGGCTGGCAAAATGGGAACC 328  
QY 161 TGGCTGATTCCTAGGACAGTGGCGGACGAGGAGGAGGAGCGCGCAGCTTCTGGAGCAGA 220  
Db 329 TGGCTGCACCTTAGGAGGTTAGTCTAGTAGGAGGAGGAGGAGGAGGAGGAGGAGG 379  
QY 221 GCCGAGACGAGACGAGTCTCTGAGTGCCTTGAACGGCCCCCTGAGCCCTACCGCCCTGCGCC 280



```

Db 6432 ACCTTGGAAATCTACTATCCCAATGATAATTCAAAATGCTGTTTACCAAGTTAGG 6491
Qy 2848 TGTGAAGGAAGTAGAGGTGGGCTTCAGGTCTCAAGGCTTCCTTAACACCCCTCT 2907
Db 6492 TGTGAAGGAAGTAGAGGTGGGCTTCAGGTCTCAAGGCTTCCTTAACACCCCTCT 6551
Qy 2908 TCCTTGGCCACCTGTTTCCGCCCTTCCACTCCCTCTACTCTCTCTAGGACTGG 2967
Db 6552 TCCTTGGCCACCTGTTTCCGCCCTTCCACTCCCTCTACTCTCTCTAGGACTGG 6611
Qy 2968 CTGATGAAGCACTGCCCAAAATTTCCCTACCCCACTTTCCCTACCCCACTTTTC 3027
Db 6612 CTGATGAAGCACTGCCCAAAATTTCCCTACCCCACTTTCCCTACCCCACTTTTC 6671
Qy 3028 CCACCAAGCTCCCAACCTGTTTGGAGCTACTGAGGAGCTTCCACTCCCTCTACTCTCTAGGACTGG 3087
Db 6672 CCACCAAGCTCCCAACCTGTTTGGAGCTACTGAGGAGCTTCCACTCCCTCTACTCTCTAGGACTGG 6731
Qy 3088 TCCCAAGCCTTTGTCATCTCAGCCCCCAGAGTATATCTGTGCTGGGAAATCTCACACA 3147
Db 6732 TCCCAAGCCTTTGTCATCTCAGCCCCCAGAGTATATCTGTGCTGGGAAATCTCACACA 6791
Qy 3148 GAACCTCAGAGACACCCCTGCTGAGCTTAAGGAGGCTTATCTCTCAGGGGGGTTTA 3207
Db 6792 GAACCTCAGAGACACCCCTGCTGAGCTTAAGGAGGCTTATCTCTCAGGGGGGTTTA 6851
Qy 3208 AGTGGCGTTTGCATATGTCGCTTATTATTTAGCGGGTGAATATTTTATCTGTA 3267
Db 6852 AGTGGCGTTTGCATATGTCGCTTATTATTTAGCGGGTGAATATTTTATCTGTA 6911
Qy 3268 GTGAGCAATCAGAGTATATGTTATGTCACAAAATTAAGGCTTTCTTATATGTTAA 3327
Db 6912 GTGAGCAATCAGAGTATATGTTATGTCACAAAATTAAGGCTTTCTTATATGTTAA 6971
Qy 3328 AAAAA 3332
Db 6972 AAAAA 6976

RESULT 21
AC096533
LOCUS Homo sapiens chromosome 1 clone RP11-6B6, WORKING DRAFT SEQUENCE, 7
DEFINITION
AC096533 AL365261
VERSION AC096533.1 GI:15638681
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEPIN.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J. and Haugen,E.D.
Direct Submission
Unpublished
2 (bases 1 to 157988)
Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J. and Haugen,E.D.
Direct Submission
Submitted (18-SEP-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Sep 18, 2001 this sequence version replaced gi:9931838.
----- Genome Center
Center: University of Washington Genome Center
Center code: UNGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
Drafting Center: SC
----- Project Information
Center project name: chr-1
Center clone name: RP11-6B6 (sc0601)
----- Summary Statistics
Sequencing vector: plasmid; L08752; 100% of reads

```

Chemistry: Dye-terminator ET; 57% of reads  
 Assembly: Dye-terminator Big Dye; 43% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 154688 bases at least Q40  
 Consensus quality: 156347 bases at least Q30  
 Consensus quality: 157058 bases at least Q20  
 Insert size: 157388; sum-of-contigs  
 Quality coverage: 7.3x in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 7 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 6264: contig of 6264 bp in length  
 6265 6364: gap of unknown length  
 6365 15785: contig of 9421 bp in length  
 15786 15885: gap of unknown length  
 15886 27947: contig of 12062 bp in length  
 27948 28048: gap of unknown length  
 28049 51992: contig of 23945 bp in length  
 51993 81154: contig of 29062 bp in length  
 81155 81254: gap of unknown length  
 81255 118813: contig of 37559 bp in length  
 118814 118913: gap of unknown length  
 118914 157988: contig of 39075 bp in length.

## FEATURES

source

Location/Qualifiers  
 1..157988  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /clone\_lib="RP11-6B6"  
 /clone="RP11-6B6"  
 1..6264  
 /note="assembly\_name:Contig22"  
 6365..15785  
 /note="assembly\_name:Contig23"  
 15886..27947  
 /note="assembly\_name:Contig24"  
 28048..51992  
 /note="assembly\_name:Contig25"  
 52093..81154  
 /note="assembly\_name:Contig26"  
 81255..118813  
 /note="assembly\_name:Contig27"  
 118914..157988  
 /note="assembly\_name:Contig28"  
 129829..129830  
 /note="gi154843|gb|J01829|TRN101S13 excised from sc0601  
 Contig28 (10917..12245)

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

QY	3028	CCACCAGCTCCAAACCCCTGTTGGAGCTACTCAGGACACAGAACGACAAAGTGCAGTT	3087	1710	AGCCACCAGGAGGAGGCTGTTCCGGCGGGGATCTGCTGACCTGCGCATCTGCTGG	1769
Db	6672	CCACCAGCTCCAAACCCCTGTTGGAGCTACTCAGGACACAGAACGACAAAGTGCAGTT	6731	5352	AGCCACCAGGAGGAGGCTGTTCCGGCGGGGATCTGCTGACCTGCGCATCTGCTGG	5411
QY	3088	TCCCAAGCCTTGTGCTCCTCAGCCCCCAGAGTATATCTGCTGTTGGGGAATCTCACACA	3147	1770	ATAGTCCCTTCTGCTGCTCCAGGTCGCCCATCTGTTTATGGCTCCATTTGTCACAG	1829
Db	6732	TCCCAAGCCTTGTGCTCCTCAGCCCCCAGAGTATATCTGCTGTTGGGGAATCTCACACA	6791	5412	ATAGTCCCTTCTGCTGCTCCAGGTCGCCCATCTGTTTATGGCTCCATTTGTCACAG	5471
QY	3148	GAACCTCAGGAGCAGCCCTGCTCAGCTAGGAGGCTTATCTCTCAGGGGGGTTA	3207	1830	TACAGCAGTCTGCTCAGCTGCTATATGCTGCTGCCGAGGCTGCTGCTGCTGCTGCT	1889
Db	6792	GAACCTCAGGAGCAGCCCTGCTCAGCTAGGAGGCTTATCTCTCAGGGGGGTTA	6851	5472	TACAGCAGTCTGCTCAGCTGCTATATGCTGCTGCCGAGGCTGCTGCTGCTGCTGCT	5531
QY	3208	AGTCCGCTTGCATAATGCTGCTCTATTTATTTAGCGGGGTGAATATTTTATCTGTA	3267	1890	ACTTTGCTACACAGGTAGTATTTGACAGAGGACTTTGGCCAAATACTCAGCGTAGAAA	1949
Db	6852	AGTCCGCTTGCATAATGCTGCTCTATTTATTTAGCGGGGTGAATATTTTATCTGTA	6911	5532	ACTTTGCTACACAGGTAGTATTTGACAGAGGACTTTGGCCAAATACTCAGCGTAGAAA	5591
QY	3268	GTGACCAATCAGAGTATATGTTATGTTGATGACAAATTAAGGCTTCTTATATGTTAA	3327	1950	CTTCCAGCAGTATGAGGCTGCTGCTCAGTGGTCCAGCTGCCGCTCCCTGTTA	2009
Db	6912	GTGACCAATCAGAGTATATGTTATGTTGATGACAAATTAAGGCTTCTTATATGTTAA	6971	5592	CTTCCAGCAGTATGAGGCTGCTGCTCAGTGGTCCAGCTGCCGCTCCCTGTTA	5651
QY	3328	AAAAA 3332		2010	GCCCCATGGGCTGCCGGCTGCCGCCAGTTCTTCTGCTGCCAAAGTAATGTCGCTCT	2069
Db	6972	AAAAA 6976		5652	GCCCCATGGGCTGCCGGCTGCCGCCAGTTCTTCTGCTGCCAAAGTAATGTCGCTCT	5711
RESULT 20				2070	CTGCTGCCACCTTGTGCTGCTGAGGTGCTGCTGACAGCTGGGGGCTGCCCT	2129
AX267731				5712	CTGCTGCCACCTTGTGCTGCTGAGGTGCTGCTGACAGCTGGGGGCTGCCCT	5771
LOCUS	AX267731	6976 bp	DNA	linear	PAT 26-OCT-2001	
DEFINITION	Sequence 705 from Patent WO0173032.			2130	CTCTCTCTCCCGCTCTAGGGCTGCTGAGGCTTCCAGAGGCTTCCAGGGGTTTCAGTTC	2189
ACCESSION	AX267731			5772	CTCTCTCTCCCGCTCTAGGGCTGCTGAGGCTTCCAGAGGCTTCCAGGGGTTTCAGTTC	5831
VERSION	AX267731.1	GI:16516403		2190	TGGACTTATACAGGGAGGCGCAGAGGCTCCATGCTGCAATGGAATGCGGGGACTCTGAGGT	2249
KEYWORDS	human.			5832	TGGACTTATACAGGGAGGCGCAGAGGCTCCATGCTGCAATGGAATGCGGGGACTCTGAGGT	5891
SOURCE	human.			2250	GGATTACCCAGGCTCAGGGTTAACTAGCTAGCTAGCTAGTTGAGACACACCTAGAGAGGS	2309
ORGANISM	Homo sapiens			5892	GGATTACCCAGGCTCAGGGTTAACTAGCTAGCTAGCTAGTTGAGACACACCTAGAGAGGS	5951
REFERENCE				2310	TTTTTGGAGCTGAATAAATCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2369
AUTHORS	Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y., Kados, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H., Vedwick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A., Hepler, W.T. and Henderson, R.A.			5952	TTTTTGGAGCTGAATAAATCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6011
TITLE	Compositions and methods for the therapy and diagnosis of prostate cancer.			2370	AGCTTCGTTTAAATGCTAGCTCTTGCATGGGAGTTTCTAGGATGAACACTCTCCATGGGA	2429
JOURNAL	Patent: WO 0173032-A 705 04-OCT-2001; CORIXA CORPORATION (US)			6012	AGCTTCGTTTAAATGCTAGCTCTTGCATGGGAGTTTCTAGGATGAACACTCTCCATGGGA	6071
FEATURES	Location/Qualifiers			2430	TTTGAACATATGAC--TTATTTTGTAGGGGAAGAGTCTCTGAGGGGCAACACACAGAACA	2487
source	1..6976			6072	TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCTCTGAGGGGCAACACACAGAACA	6131
BASE COUNT	1312 a 1996 c 1961 g 1706 t		1 others	2488	GGTCCCTCAGCCACAGCAGTCTGCTTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2547
ORIGIN				6132	GGTCCCTCAGCCACAGCAGTCTGCTTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6191
Query Match	53.2%; Score 1815.8; DB 6; Length 6976;			2548	GGATGTCGCTGTTGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2607
Best Local Similarity	.98.7%; Pred. No. 2e-300;			6192	GGATGTCGCTGTTGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6251
Matches 1841; Conservative	0; Mismatches 22; Indels 2; Gaps 1;			2608	CTTATTTTAAACAAAGTAGAGGGAATCCATTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCT	2667
QY	1470	CCTACACTGGCTCCCTCTACACCGGGGAGAGAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1529	6252	CTTATTTTAAACAAAGTAGAGGGAATCCATTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCT	6311
Db	5112	CCTGCTCTCTCCCTTTTCTTCCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5171	2668	TATTTGGTGGGTTGGGATCCCAACAAATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2727
QY	1530	GGGACACTGGAGTGTACAGTGGAGGAGGCTGATGACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1589	6312	TATTTGGTGGGTTGGGATCCCAACAAATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6371
Db	5172	GGGACACTGGAGTGTACAGTGGAGGAGGCTGATGACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5231	2728	CTGATCATTTGCCAGAAATCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2787
QY	1590	AGCTGGAGTCCCTTCCCTTAATGACACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1649	6372	CTGATCATTTGCCAGAAATCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	6431
Db	5232	AGCTGGAGTCCCTTCCCTTAATGACACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5291	2788	ACCTTGGAAATTTCTACTCATCCCAAAATGATAATTTCCAAATGCTGTTTACCACCAAGTTAGGG	2847
QY	1650	CTCACCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1709			
Db	5292	CTCACCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	5351			

QY 3215 TTTCAGTAATAGTCTCTTATTTATTTAGCGGGGTGAATATTTTATCTACTAGTAGCA 3274  
|||||  
Db  
2337 TTTCAGTAATAGTCTCTTATTTATTTAGCGGGGTGAATATTTTATCTACTAGTAGCA 2396  
|||||  
QY 3275 ATCAGAGTATAATGTTATGTTGACAAATTAAGGCTTCTTATATGTTTAAAAA 3334  
|||||  
Db  
2397 ATCAGAGTATAATGTTATGTTGACAAATTAAGGCTTCTTATATGTTTAAAAA 2456  
|||||  
QY 3335 AAAAAA 3355  
|||||  
Db  
2457 AAAAAA 2477  
|||||  
RESULT 19  
AX200996 6976 bp DNA linear PAT 29-AUG-2001  
LOCUS  
DEFINITION Sequence 626 from Patent WO0151633.  
ACCESSION AX200996  
VERSION AX200996.1 GI:15390823  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 6976)  
Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y.,  
Read, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W.,  
Stolk, J.A., Skeiky, Y.A., Wang, A. and Meagher, M.J.  
Compositions and methods for the therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0151633-A 626 19-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers  
source 1..6976  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1312 a 1996 c 1961 g 1706 t 1 others  
ORIGIN  
Query Match 53.2%; Score 1815.8; DB 6; Length 6976;  
Best Local Similarity 98.7%; Pred. No. 2e-300;  
Matches 1841; Conservative 0; Mismatches 22; Indels 2; Gaps 1;  
QY 1470 CTTACACACTGGGCTCCCTTACACCGGAGAGAGAGTGTTCCTGCCCAATACCGAG 1529  
|||||  
Db  
5112 CCGTCTCTCCCTTTCTTCCCTTCTGCTTACCTTACCTTACCTTACCTTACCTTACCT 5171  
|||||  
QY 1530 GGGACACTGGAGTGTAGCAGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1589  
|||||  
Db  
5172 GGGACACTGGAGTGTAGCAGTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5231  
|||||  
QY 1590 AGCTGGAGCTCCCTTCCCTTAATGGACAGCTGGGTGCTGGAGGAGTGGCTGCTCCAC 1649  
|||||  
Db  
5232 AGCTGGAGCTCCCTTCCCTTAATGGACAGCTGGGTGCTGGAGGAGTGGCTGCTCCAC 5291  
|||||  
QY 1650 CTCACACCGGCTCTCGGGGGCTCTGCTGCTGATGTCCTGACGAGTGGTGGTGGG 1709  
|||||  
Db  
5292 CTCACACCGGCTCTCGGGGGCTCTGCTGCTGATGTCCTGACGAGTGGTGGTGGG 5351  
|||||  
QY 1710 AGCCACCGAGGAGAGTGGTTCGGGGCGGGGAGTGGTGGAGTGGTGGTGGTGGTGG 1769  
|||||  
Db  
5352 AGCCACCGAGGAGAGTGGTTCGGGGCGGGGAGTGGTGGAGTGGTGGTGGTGGTGG 5411  
|||||  
QY 1770 ATAGTGCTTCTCTGCTGCCAGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1829  
|||||  
Db  
5412 ATAGTGCTTCTCTGCTGCCAGTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 5471  
|||||  
QY 1830 TCAGCAGTGTCTACAGTGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1889  
|||||  
Db  
5472 TCAGCAGTGTCTACAGTGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5531  
|||||  
QY 1890 ACTTGTCTACACAGTGTATTTGACAGAGGAGTGGGCGGCGGCTGCTGCTGCTGCTGCT 1949  
|||||  
5532 ACTTTGCTACACAGTGTATTTGACAGAGGAGTGGGCGGCGGCTGCTGCTGCTGCTGCT 5591  
1950 CTTTCCAGACATTTGGGGTGGAGGCGCTGCTCCTCAGTGGTCCAGTCCCGCTCTCTGTTA 2009  
5592 CTTTCCAGACATTTGGGGTGGAGGCGCTGCTCCTCAGTGGTCCAGTCCCGCTCTCTGTTA 5651  
2010 GCGCCATGGGCTGCGGGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069  
5652 GCGCCATGGGCTGCGGGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5711  
2070 CTGCTGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2129  
5712 CTGCTGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5771  
2130 CTGCTCTCTCCAGTCTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2189  
5772 CTGCTCTCTCCAGTCTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5831  
2190 TGGACTTATACAGGAGGCGCAGAGGCTTCCATGCACTGCAATGCGGGGAGTCTGCAGGT 2249  
5832 TGGACTTATACAGGAGGCGCAGAGGCTTCCATGCACTGCAATGCGGGGAGTCTGCAGGT 5891  
2250 GGATTAACCGAGCTCAGGCTTAACAGCTAGCTTACCTCTCTAGTGGAGGCTTCCAGGGG 2309  
5892 GGATTAACCGAGCTCAGGCTTAACAGCTAGCTTACCTCTCTAGTGGAGGCTTCCAGGGG 5951  
2310 TTTTGGGAGCTCAATAACTCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2369  
5952 TTTTGGGAGCTCAATAACTCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6011  
2370 AGCTTCTGTTAATGAGTCTTTCGATGGAGTTCCTAGGATGAACACACTCTCTCCATGGGA 2429  
6012 AGCTTCTGTTAATGAGTCTTTCGATGGAGTTCCTAGGATGAACACACTCTCTCCATGGGA 6071  
2430 TTTGAACATATGAC--TTATTTGTTAGGGAAGAGTCTGAGGGGCAACACACAGAACCA 2487  
6072 TTTGAACATATGAAAGTTATTTGTTAGGGAAGAGTCTGAGGGGCAACACACAGAACCA 6131  
2488 GTTCCCTCAGCCACAGCAGTCTTTTGTGCTGATCCACCCCTCTACCTTTATCA 2547  
6132 GTTCCCTCAGCCACAGCAGTCTTTTGTGCTGATCCACCCCTCTACCTTTATCA 6191  
2548 GGATGCGCTGTTGGTCTCTTGTGCTATCACAGACACAGGCAATTTAAATATTTAA 2607  
6192 GGATGCGCTGTTGGTCTCTTGTGCTATCACAGACACAGGCAATTTAAATATTTAA 6251  
2608 CTTATTTATTTAAACAAAGTAGAAGGAATCCATGCTAGCTTTTCTGTTGGTCTCAA 2667  
6252 CTTATTTATTTAAACAAAGTAGAAGGAATCCATGCTAGCTTTTCTGTTGGTCTCAA 6311  
2668 TATTTTGGGTAGGTTGGGATCCCAACAATCAGTCCCTGAGATAGTGTGCTATTGGG 2727  
6312 TATTTTGGGTAGGTTGGGATCCCAACAATCAGTCCCTGAGATAGTGTGCTATTGGG 6371  
2728 CTGATCATTTGCCAAGTCTTCTTCTCTGGGCTCTGGGCGGCGGCGGCGGCGGCGGCGG 2787  
6372 CTGATCATTTGCCAAGTCTTCTTCTCTGGGCTCTGGGCGGCGGCGGCGGCGGCGGCGG 6431  
2788 ACCTTGGAAATTTCTACTCATCCCAATGATTAATCCAAATGCTGTATCCCAAGTTAGG 2847  
6432 ACCTTGGAAATTTCTACTCATCCCAATGATTAATCCAAATGCTGTATCCCAAGTTAGG 6491  
2848 TGTTCAGGAAGTAGAGGTTGGGCTTTCAGTCTCAACGGCTTCCCTCAACACCCCTCT 2907  
6492 TGTTCAGGAAGTAGAGGTTGGGCTTTCAGTCTCAACGGCTTCCCTCAACACCCCTCT 6551  
2908 TCTCTTGGCCAGCTGTTTCCCCCCTTCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCT 2967  
6552 TCTCTTGGCCAGCTGTTTCCCCCCTTCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCT 6611  
2968 CTGATGAAGGCACTGCCCAAAATTTCCCTTCCCTTCCCTTCTTCTTCTTCTTCTTCTTCT 3027  
6612 CTGATGAAGGCACTGCCCAAAATTTCCCTTCCCTTCCCTTCTTCTTCTTCTTCTTCTTCT 6671

```
/clone="DKFp666D0110"
/tissue.type="stomach"
/clone.lib="666 (synonym: hsto2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
polyA_signal 2425..2430
polyA_site 2447
BASE COUNT 505 a 713 c 634 g 625 t
ORIGIN
```

Query Match 61.7%; Score 2103.8; DB 9; Length 2477;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2117; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY	1237	TGAGGGGTTCCGATGGGAGCTGGGGTGTTCCTGCAGTGGCCATCTCCCTGCTCTT	1296
DB	357	TGAGGGGTTCCGATGGGAGCTGGGGTGTTCCTGCAGTGGCCATCTCCCTGCTCTT	416
QY	1297	CTCTCTGTCATGGACCGGCTGTGTCAGCGATTTCGGCAGTGCAGTCTATTGCGCCAG	1356
DB	417	CTCTCTGTCATGGACCGGCTGTGTCAGCGATTTCGGCAGTGCAGTCTATTGCGCCAG	476
QY	1357	TGTGGCAGCTTTCCTCTGTGCTGCCGCTGCCATGCTGTCCACAGTGTGGCGTGGT	1416
DB	477	TGTGGCAGCTTTCCTCTGTGCTGCCGCTGCCATGCTGTCCACAGTGTGGCGTGGT	536
QY	1417	GACAGCTTCAGCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCTGCCCTACAC	1476
DB	537	GACAGCTTCAGCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCTGCCCTACAC	596
QY	1477	ACTGGCTTCCCTTACACCGGGAGAGCAGGTGTTCCTGCCCAATACAGGGGACAC	1536
DB	597	ACTGGCTTCCCTTACACCGGGAGAGCAGGTGTTCCTGCCCAATACAGGGGACAC	656
QY	1537	TGGAGGTGTAGCAGTGCAGCAGCCTGATGACAGCTTCTGCCAGGCCCTAAGCCTGG	1596
DB	657	TGGAGGTGTAGCAGTGCAGCAGCCTGATGACAGCTTCTGCCAGGCCCTAAGCCTGG	716
QY	1597	AGCTCCCTTCCCTAATGACACGTGGGTGTGAGGAGCAGTGGCTGTGCCACCTCCACC	1656
DB	717	AGCTCCCTTCCCTAATGACACGTGGGTGTGAGGAGCAGTGGCTGTGCCACCTCCACC	776
QY	1657	CGCGCTGTGGGGGCTCTGCCTGTGATGTCTCGGTACGTGTGGTGGGTGAGCCAC	1716
DB	777	CGCGCTGTGGGGGCTCTGCCTGTGATGTCTCGGTACGTGTGGTGGGTGAGCCAC	836
QY	1717	CGAGCCAGGTGGTTCGGGCGGGGATCTGCTGGACCTGCCATCTCGGATAGTGC	1776
DB	837	CGAGCCAGGTGGTTCGGGCGGGGATCTGCTGGACCTGCCATCTCGGATAGTGC	896
QY	1777	CTTCTGCTGTCCAGGTGGCCCATCCCTGTTTATGGGCTCCATTTGTCCAGCTCAGCCA	1836
DB	897	CTTCTGCTGTCCAGGTGGCCCATCCCTGTTTATGGGCTCCATTTGTCCAGCTCAGCCA	956
QY	1837	GTCTGTACTGCTATATGTTGTCTGGGAGGCTGGGTGTGGTCTGCCCATTTACTTTGC	1896
DB	957	GTCTGTACTGCTATATGTTGTCTGGGAGGCTGGGTGTGGTCTGCCCATTTACTTTGC	1016
QY	1897	TACACAGTGTATTGTGACAGAGCGACTTGGCCAAATACCTACGCTAGAAAAGTCCAG	1956
DB	1017	TACACAGTGTATTGTGACAGAGCGACTTGGCCAAATACCTACGCTAGAAAAGTCCAG	1076
QY	1957	CACATTGGGGTGGAGGGCTGCCCTACGTGGTCCAGCTCCCGCTCCTGTAGCCCAT	2016
DB	1077	CACATTGGGGTGGAGGGCTGCCCTACGTGGTCCAGCTCCCGCTCCTGTAGCCCAT	1136
QY	2017	GGGGTGGCGGGCTGGCGCCAGTTTCTGTTGCTGCCAAATATGTGGCTCTCTGTGC	2076
DB	1137	GGGGTGGCGGGCTGGCGCCAGTTTCTGTTGCTGCCAAATATGTGGCTCTCTGTGC	1196
QY	2077	CACCTGTGCTGTGCTGAGTGTGCTGACAGTGTGGGGCTGGGGCTGCCCTCTCTCT	2136
DB	1197	CACCTGTGCTGTGCTGAGTGTGCTGACAGTGTGGGGCTGGGGCTGCCCTCTCTCT	1256

QY	2137	CTCCCCAGTCTCTAGGGCTGCTGACTGGAGGCTTCCAAAGGGGTTTTCAGTCTGGACTT	2196
DB	1257	CTCCCCAGTCTCTAGGGCTGCTGACTGGAGGCTTCCAAAGGGGTTTTCAGTCTGGACTT	1316
QY	2197	ATACAGGAGGCCAGAGGGCTCCATGACTGGAATGCGGGGACTCTGCAGGTGGATTAC	2256
DB	1317	ATACAGGAGGCCAGAGGGCTCCATGACTGGAATGCGGGGACTCTGCAGGTGGATTAC	1376
QY	2257	CCAGCTCAGGGTTAAACAGCTAGCTCTCTAGTTCAGACACACCTAGAGAGGGTTTTTGG	2316
DB	1377	CCAGCTCAGGGTTAAACAGCTAGCTCTCTAGTTCAGACACACCTAGAGAGGGTTTTTGG	1436
QY	2317	GAGCTGAATAAAGTACAGTACCTGTTTCCCATCTCTAAAGCCCTTAACCTGACAGTTGG	2376
DB	1437	GAGCTGAATAAAGTACAGTACCTGTTTCCCATCTCTAAAGCCCTTAACCTGACAGTTGG	1496
QY	2377	TTTAATGTAGCTCTTTCATGGAGTTCCTAGGATGAACACCTCTCCATGGATTGAAAC	2436
DB	1497	TTTAATGTAGCTCTTTCATGGAGTTCCTAGGATGAACACCTCTCCATGGATTGAAAC	1556
QY	2437	ATATG--ACTTATTTGTAGGGAGAGTCTCTGAGGGCAACACACAGAACAGGTCGCC	2494
DB	1557	ATATGAAAGTTATTTGTAGGGAGAGTCTCTGAGGGCAACACACAGAACAGGTCGCC	1616
QY	2495	TCAGCCACAGCAGTCTTTTGTGTCATCCACCCCTCTTACCTTTTATCAGGATGTG	2554
DB	1617	TCAGCCACAGCAGTCTTTTGTGTCATCCACCCCTCTTACCTTTTATCAGGATGTG	1676
QY	2555	GCCTGTGGTCTCTCTGTTGCCATCAGAGACACAGGCAATTTAAATATTTAACTTATTT	2614
DB	1677	GCCTGTGGTCTCTCTGTTGCCATCAGAGACACAGGCAATTTAAATATTTAACTTATTT	1736
QY	2615	ATTTAACAAAGTAGAAGGAATCCATTTCTAGCTTTTCTGTGTGGTGTCTAAATTTGG	2674
DB	1737	ATTTAACAAAGTAGAAGGAATCCATTTCTAGCTTTTCTGTGTGGTGTCTAAATTTGG	1796
QY	2675	GTAGGTGGGGGATCCCAACAAATCAGTCCCTCAGATAGCTGGTGTCTATGGGCTGATCA	2734
DB	1797	GTAGGTGGGGGATCCCAACAAATCAGTCCCTCAGATAGCTGGTGTCTATGGGCTGATCA	1856
QY	2735	TTGCCAGAAATCTTCTCTCTGGGCTTGGCCCCCAAAATGCCTTAACCCAGGACCTTGG	2794
DB	1857	TTGCCAGAAATCTTCTCTCTGGGCTTGGCCCCCAAAATGCCTTAACCCAGGACCTTGG	1916
QY	2795	AAATTTACTCATCCCAAAATATATCCAAATGTGTTTACCAAGTTTAGGGTGTGAA	2854
DB	1917	AAATTTACTCATCCCAAAATATATCCAAATGTGTTTACCAAGTTTAGGGTGTGAA	1976
QY	2855	GGAAAGTAGAGGTGGGGCTTCAGGTCCTCAACGGCTTCCCTAACCCACCTCTCTCTTG	2914
DB	1977	GGAAAGTAGAGGTGGGGCTTCAGGTCCTCAACGGCTTCCCTAACCCACCTCTCTCTTG	2036
QY	2915	GCCAGCCTGGTTTCCCCCACTTCCACTCCCTCTACTCTCTTAGGACTGGGCTGATGA	2974
DB	2037	GCCAGCCTGGTTTCCCCCACTTCCACTCCCTCTACTCTCTTAGGACTGGGCTGATGA	2096
QY	2975	AGGACTGCCCAAAATTTCCCTACCCCACTTTCCCTACCCCACTTTCCCACTTTCCCA	3034
DB	2097	AGGACTGCCCAAAATTTCCCTACCCCACTTTCCCTACCCCACTTTCCCACTTTCCCA	2156
QY	3035	GTCCCAACACCTGTTTGGAGTCTGAGGACCAAGCAAGCAAGTGGGTTTCCCAAG	3094
DB	2157	GTCCCAACACCTGTTTGGAGTCTGAGGACCAAGCAAGCAAGTGGGTTTCCCAAG	2216
QY	3095	CTTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGTTTGGGGAATCTCACAGAAATCT	3154
DB	2217	CTTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGTTTGGGGAATCTCACAGAAATCT	2276
QY	3155	AGGAGACCCCTGCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAAAGTCCG	3214
DB	2277	AGGAGACCCCTGCTGAGCTAAGGAGGTCTTATCTCTCAGGGGGGTTTAAAGTCCG	2336







```
Db 781 CACATTGGGGTGGAGGGGCTCCCTCACTGGGTCCAGCTCCCGCTCCTGTAGCCCAT 840
Qy 2017 GGGGCTGCCGGCTGGCGCCAGTTCTGTGTCGCCAAAGTATGCTCTCTGCTGC 2076
Db 841 GGGGCTGCCGGCTGGCGCCAGTTCTGTGTCGCCAAAGTATGCTCTCTGCTGC 900
Qy 2077 CACCTGTGCTGCTGAGTGCGTGTAGCTGCACAGCTGGGGGCTGGGGGCTCCCTCCTCT 2136
Db 901 CACCTGTGCTGCTGAGTGCGTGTAGCTGCACAGCTGGGGGCTGGGGGCTCCCTCCTCT 960
Qy 2137 CTCCTCCAGTCTCTAGGGCTGCTGACTGGAGGCGCTTCAAGGGGTTTCAGTCTGGAATT 2196
Db 961 CTCCTCCAGTCTCTAGGGCTGCTGACTGGAGGCGCTTCAAGGGGTTTCAGTCTGGAATT 1020
Qy 2197 ATACAGGAGGCGCAGAGGGCTCCATGCACTGGAATGCGGAGCTCTGCAAGTGGAATTAC 2256
Db 1021 ATACAGGAGGCGCAGAGGGCTCCATGCACTGGAATGCGGAGCTCTGCAAGTGGAATTAC 1080
Qy 2257 CCAGGCTCAGGTTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAGGGTTTTGG 2316
Db 1081 CCAGGCTCAGGTTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAGGGTTTTGG 1140
Qy 2317 GAGCTGAATAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCTTAACTGCAAGCTTCG 2376
Db 1141 GAGCTGAATAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCTTAACTGCAAGCTTCG 1200
Qy 2377 TTTAATGTAGCTCTTGTGATGGAGTTCTTAGGATGAACACCTCCTCCATGGGATTTGAAC 2436
Db 1201 TTTAATGTAGCTCTTGTGATGGAGTTCTTAGGATGAACACCTCCTCCATGGGATTTGAAC 1260
Qy 2437 ATATG--ACTTATTTGTAGGGAAGAGCTCTGAGGGGCAACACACAGAGGAGGTCCTCC 2494
Db 1261 ATATGAAAGTTATTTGTAGGGAAGAGCTCTGAGGGGCAACACACAGAGGAGGTCCTCC 1320
Qy 2495 TCAGGCCACAGACACTCTTTTGTGATCAACCCCTCTTACCTTTTATCAGAGTGTG 2554
Db 1321 TCAGGCCACAGACACTCTTTTGTGATCAACCCCTCTTACCTTTTATCAGAGTGTG 1380
Qy 2555 GCCTGTGCTCTCTCTGTGTCATCACAGACACAGAGCAATTTAAATATTTAACTTATTT 2614
Db 1381 GCCTGTGCTCTCTGTGTCATCACAGACACAGAGCAATTTAAATATTTAACTTATTT 1440
Qy 2615 ATTTAAACAGTAGAAGGAATCCATGCTAGCTTTTCTGTGTTGCTCTAAATATTTGG 2674
Db 1441 ATTTAAACAGTAGAAGGAATCCATGCTAGCTTTTCTGTGTTGCTCTAAATATTTGG 1500
Qy 2675 GTAGGGTGGGATCCCAACAACTCAGTCCCTCAGATAGCTGTGCTATTTGGCTGTATCA 2734
Db 1501 GTAGGGTGGGATCCCAACAACTCAGTCCCTCAGATAGCTGTGCTATTTGGCTGTATCA 1560
Qy 2735 TTGCCAGAACTCTTCTCTCTGGGCTCTGGCCCTCCCAAAATGCCCTAACCCAGGACCTTGG 2794
Db 1561 TTGCCAGAACTCTTCTCTCTGGGCTCTGGCCCTCCCAAAATGCCCTAACCCAGGACCTTGG 1620
Qy 2795 AAATCTACTATCCCAATATATTCAAATGCTGTATCCCAAGTTAGGGTTGAA 2854
Db 1621 AAATCTACTATCCCAATATATTCAAATGCTGTATCCCAAGTTAGGGTTGAA 1680
Qy 2855 GGAAGTGTAGGGTGGGCTTCAAGTCTCAACGGCTTCCCTAACCAACCCCTCTCTCTCTTG 2914
Db 1681 GGAAGTGTAGGGTGGGCTTCAAGTCTCAACGGCTTCCCTAACCAACCCCTCTCTCTCTTG 1740
Qy 2915 GCCAGCCTCTGTTTCCCTCCCTTCCACTCCCTCTACTCTCTTAGGACTGGGCTGTATGA 2974
Db 1741 GCCAGCCTCTGTTTCCCTCCCTTCCACTCCCTCTACTCTCTTAGGACTGGGCTGTATGA 1800
Qy 2975 AGGCACTGCCAAAATTTCCCTTACCCCACTTTCCCTTACCCCACTTTCCCACTCA 3034
Db 1801 AGGCACTGCCAAAATTTCCCTTACCCCACTTTCCCTTACCCCACTTTCCCACTCA 1860
Qy 3035 GCTCCACACCTCTTTTGGAGCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3094
Db 1861 GCTCCACACCTCTTTTGGAGCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Qy 3095 CTTTTTCCATCTCAGCCCCCAGAGTATATCTGTCTTGGGAATCTCACACAGAACTC 3154
Db 1921 CTTTTTCCATCTCAGCCCCCAGAGTATATCTGTCTTGGGAATCTCACACAGAACTC 1980
Qy 3155 AGAGCACCCCTGCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAAGTGCG 3214
Db 1981 AGAGCACCCCTGCTGAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAAGTGCG 2049
Qy 3215 TTTGCAATATGTCGCTTATTTATTTAGCGGGTGAATATTTATCTGTAAGTGAGCA 3274
Db 2041 TTTGCAATATGTCGCTTATTTATTTAGCGGGTGAATATTTATCTGTAAGTGAGCA 2100
Qy 3275 ATCAGAGTATAATGTTTATGTTGACAAAATTAAGGCTTCTTATATGTTTA 3326
Db 2101 ATCAGAGTATAATGTTTATGTTGACAAAATTAAGGCTTCTTATATGTTTA 2152

RESULT 17
AR112294
LOCUS AR112294 2143 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 15 from patent US 6130043.
ACCESSION AR112294
VERSION AR112294.1 GI:14092194
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2143)
AUTHORS Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N.,
Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D.,
Roberts-Rapp, L., Russell, J.C., Stroupe, S.D. and Yu, H.
Reagents and methods useful for detecting diseases of the prostate
Patent: US 6130043-A 15 10-OCT-2000;
JOURNAL Location/Qualifiers
FEATURES
source 1..2143
BASE COUNT 418 a 621 c 563 g 541 t
ORIGIN
Query Match 62.0%; Score 2114.8; DB 6; Length 2143;
Best Local Similarity 99.88; Pred. No. 0;
Matches 2139; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

Qy 1185 ACCAGGCGCTGCCAGAGCTGAGCGGGCACCGGGACCGGAGACACTATGATGAAGCG 1244
Db 1 ACCAGGCGCTGCCAGAGCTGAGCGGGCACCGGGACCGGAGACACTATGATGAAGCG 60
Qy 1245 TTCGGATGGCGAGCTGGGGCTGTCTCTGAGTGGCGCATCTCCCTGCTCTCTCTCTG 1304
Db 61 TTCGGATGGCGAGCTGGGGCTGTCTCTGAGTGGCGCATCTCCCTGCTCTCTCTCTG 120
Qy 1305 TCATGACCGGCTGGTTCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 1364
Db 121 TCATGACCGGCTGGTTCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 180
Qy 1365 CTTTCCCTGTGGTGGCGGTCACATGCTGCTCCACAGTGTGGCGGTGGTGCACAGCTT 1424
Db 181 CTTTCCCTGTGGTGGCGGTCACATGCTGCTCCACAGTGTGGCGGTGGTGCACAGCTT 240
Qy 1425 CAGCCGCTCAGCGGCTTCACTTCTCAGCCCTGCAGATCTCCCTACACACTGGCCT 1484
Db 241 CAGCCGCTCAGCGGCTTCACTTCTCAGCCCTGCAGATCTCCCTACACACTGGCCT 300
Qy 1485 CCCTCTACACCGGAGAGCAGGTGTCTCTGCGCAAAATACCGAGGGGACACTGGAGGTG 1544
Db 301 CCCTCTACACCGGAGAGCAGGTGTCTCTGCGCAAAATACCGAGGGGACACTGGAGGTG 360
Qy 1545 CTAGCAGTAGGACAGCTGATGACAGCTTCTGCGCAGGCGCTTAAAGCTGGAGCTCCCT 1604
Db 361 CTAGCAGTAGGACAGCTGATGACAGCTTCTGCGCAGGCGCTTAAAGCTGGAGCTCCCT 420
```

Db 3891 CTCAGTACCTGGTTTCCCATCTCTAAGCCCTCTAACTGACGCTTGGTTTAATGTAGCT 3950  
QY 2389 CTTGATGGGAGTTTCTAGGATGAACACTCTCCATGGGATTTGAACATATG - ACTTA 2446  
Db 3951 CTTGATGGGAGTTTCTAGGATGAACACTCTCCATGGGATTTGAACATATGAAAGTTA 4010  
QY 2447 TTTGAGGGAGAGTCTCTAGGGGCAACACAGAACAGGTCCTCCCTCAGGCCACAGC 2506  
Db 4011 TTTGAGGGAGAGTCTCTAGGGGCAACACAGAACAGGTCCTCCCTCAGGCCACAGC 4070  
QY 2507 ACTGCTCTTTTGTCTGATCCACCCCTCTACCTTTTATCAGGATGGGCTGTGGTCC 2566  
Db 4071 ACTGCTCTTTTGTCTGATCCACCCCTCTACCTTTTATCAGGATGGGCTGTGGTCC 4130  
QY 2567 TTTCTGTGCCATCACAGACACAGGCAATTAATATTTAACTTTATTTTAAACAAGT 2626  
Db 4131 TTTCTGTGCCATCACAGACACAGGCAATTAATATTTAACTTTATTTTAAACAAGT 4190  
QY 2627 AGAAGGGAATCCATTTGCTAGCTTTTCTGTGTGTCTAATATTTGGGTAGGGTGGGG 2686  
Db 4191 AGAAGGGAATCCATTTGCTAGCTTTTCTGTGTGTCTAATATTTGGGTAGGGTGGGG 4250  
QY 2687 ATCCCAACATCAGTCCCTGATAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2746  
Db 4251 ATCCCAACATCAGTCCCTGATAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4310  
QY 2747 TCTTCTCTGGGCTCTGGGCCCCCAAAATGCCTAACCCAGGAGCTTTGGAATTTCTACTCA 2806  
Db 4311 TCTTCTCTGGGCTCTGGGCCCCCAAAATGCCTAACCCAGGAGCTTTGGAATTTCTACTCA 4370  
QY 2807 TCCCAATGATATTTCAATGCTTTTACCAGAGTTAGGCTGTTGAAGGAAGGTAGAG 2866  
Db 4371 TCCCAATGATATTTCAATGCTTTTACCAGAGTTAGGCTGTTGAAGGAAGGTAGAG 4430  
QY 2867 GTGGGGCTTCAAGGCTTCCCTTAACCAACCCCTCTCTCTTGGCCCCAGGCTGTGT 2926  
Db 4431 GTGGGGCTTCAAGGCTTCCCTTAACCAACCCCTCTCTCTTGGCCCCAGGCTGTGT 4490  
QY 2927 TCCCCCACTTCCACTCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986  
Db 4491 TCCCCCACTTCCACTCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550  
QY 2987 AATTTTCCCTTACCCCACTTCCCTTACCCCACTTCCCTTACCCCACTTCCCTTACCCCA 3046  
Db 4551 AATTTTCCCTTACCCCACTTCCCTTACCCCACTTCCCTTACCCCACTTCCCTTACCCCA 4610  
QY 3047 TGTGTGAGTACTGACGAGCAGAGCAAAAGTGGGTTTCCCAAGCTTTTGTGCATC 3106  
Db 4611 TGTGTGAGTACTGACGAGCAGAGCAAAAGTGGGTTTCCCAAGCTTTTGTGCATC 4670  
QY 3107 TCAGCCCCAGATATCTGTGCTGGGGAATCTCACAGAAACTCAGGAGCACCCCC 3166  
Db 4671 TCAGCCCCAGATATCTGTCTGGGGAATCTCACAGAAACTCAGGAGCACCCCC 4730  
QY 3167 TGCTGAGTAAAGGAGTCTTATCTCTCAGGGGGGTTTAAAGTGGGCTTTGCAATATG 3226  
Db 4731 TGCTGAGTAAAGGAGTCTTATCTCTCAGGGGGGTTTAAAGTGGGCTTTGCAATATG 4790  
QY 3227 TCCTCTTATTTATTTAGCGGGTGAATATTTTATCTGTAAGTGAGCAATCAGAGTATAA 3286  
Db 4791 TCCTCTTATTTATTTAGCGGGTGAATATTTTATCTGTAAGTGAGCAATCAGAGTATAA 4850  
QY 3287 TGTTTATGGTACAAAATTAAGGCTTTCTTATATCTTTTAAAA 3330  
Db 4851 TGTTTATGGTACAAAATTAAGGCTTTCTTATATCTTTTAAAAA 4894

RESULT 16

ARL12295

LOCUS

DEFINITION Sequence 16 from patent US 6130043.

ACCESSION ARL12295

VERSION ARL12295.1 GI:14092195

linear PAT 16-MAY-2001

## KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2152)

AUTHORS Billing-Medel, P.A., Cohen, M., Colpitts, T.L., Friedman, P.N.,

Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D.,

Roberts-Rapp, L., Russell, J.C., Stroupe, S.D., and Yu, H.

TITLE Reagents and methods useful for detecting diseases of the prostate

JOURNAL Patent: US 6130043-A 16 10-OCT-2000;

FEATURES Location/Qualifiers

source 1..2152

BASE COUNT 419 a 622 c 569 g 542 t

ORIGIN /organism="unknown"

Query Match 62.7%; Score 2136.4; DB 6; Length 2152;

Best local similarity 99.9%; Pred. No. 0;

Matches 2149; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1177 GGGGCTGTACAGGGGCTGCCAGAGCTAGCGGGCCAGCGGGCCAGAGACTATGA 1236  
Db 1 GGGGCTGTACAGGGGCTGCCAGAGCTAGCGGGCCAGCGGGCCAGAGACTATGA 60  
QY 1237 TGAAGCGCTTCGATGGCAGCGCTGGGGCTGTTCCTGCAGTGCAGCATCTCCTGTGCTT 1296  
Db 61 TGAAGCGCTTCGATGGGAGCGCTGGGGCTGTTCCTGCAGTGCAGCATCTCCTGTGCTT 120  
QY 1297 CTCCTGTGTCATGGACCGGCTGGTGAGCGATTCGGCAGTTCGAGCAGTCTATTTGGCCAG 1356  
Db 121 CTCCTGTGTCATGGACCGGCTGGTGAGCGATTCGGCAGTTCGAGCAGTCTATTTGGCCAG 180  
QY 1357 TGTGGCAGCTTTCCTGTGGCTGCCGCTGCCAGTTCACCTTCTCAGCCCTCGAGATCTCCTGCCATAC 1416  
Db 181 TGTGGCAGCTTTCCTGTGGCTGCCGCTGCCAGTTCACCTTCTCAGCCCTCGAGATCTCCTGCCATAC 240  
QY 1417 GACAGCTTCAGCGCCCTCACCGGTTTCACTTCTCAGCCCTCGAGATCTCCTGCCATAC 1476  
Db 241 GACAGCTTCAGCGCCCTCACCGGTTTCACTTCTCAGCCCTCGAGATCTCCTGCCATAC 300  
QY 1477 ACTGGGCTCCCTTACCACCGGGAGAGCAGGTGTCTGCCCAATACCGAGGGGACAC 1536  
Db 301 ACTGGGCTCCCTTACCACCGGGAGAGCAGGTGTCTGCCCAATACCGAGGGGACAC 360  
QY 1537 TGGAGTGTACAGTGAAGCAGCCTGATGACACAGCTTCTCCAGCGCCCTTAAGCCCTGG 1596  
Db 361 TGGAGTGTACAGTGAAGCAGCCTGATGACACAGCTTCTCCAGCGCCCTTAAGCCCTGG 420  
QY 1597 AGCTCCCTTCCCTAATGAGCAGCTGGGTCTGGAGCAGTGGCCTGCTCCCACTCCACC 1656  
Db 421 AGCTCCCTTCCCTAATGAGCAGCTGGGTCTGGAGCAGTGGCCTGCTCCCACTCCACC 480  
QY 1657 CCGGCTCTCGGGGCTCTGCTGTGATGTCTCCGTAGTGTGGTGGGTGAGCCAC 1716  
Db 481 CCGGCTCTCGGGGCTCTGCTGTGATGTCTCCGTAGTGTGGTGGGTGAGCCAC 540  
QY 1717 CGAGCCAGGCTGGTTCGGGGCGGGGCTATGCTGGACCTCGCAGCTTCCTGGATAGTGC 1776  
Db 541 CGAGCCAGGCTGGTTCGGGGCGGGGCTATGCTGGACCTCGCAGCTTCCTGGATAGTGC 600  
QY 1777 CTTCTGTGTGCCAGGTGGGCCCCATCCCTGTTTATGGGCTCCATTTGCCAGCTCAGCA 1836  
Db 601 CTTCTGTGTGCCAGGTGGGCCCCATCCCTGTTTATGGGCTCCATTTGCCAGCTCAGCA 660  
QY 1837 GTCTGTCACTGCTATATGTTGTCTGCGCGAGGCTGGTCTGGTGGCCATTTACTTTGC 1896  
Db 661 GTCTGTCACTGCTATATGTTGTCTGCGCGAGGCTGGTCTGGTGGCCATTTACTTTGC 720  
QY 1897 TACACAGTGTATTTGACAAGAGGCTTGGGCCAAATACCTCAGCGTGAAGAACTTCCAG 1956  
Db 721 TACACAGTGTATTTGACAAGAGGCTTGGGCCAAATACCTCAGCGTGAAGAACTTCCAG 780  
QY 1957 CACATTTGGGGTGGAGGGGCTGCCTCACTGGGGTCCAGCTCCCGCTCTGTTAGCCCAT 2016



3591 CTGGCGCCAGTTCTTGTGTCGCAAGTAAGTGTGCTCTCTGCTCCACCCCTGTGCTG 3650  
Db  
2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGCTCCCTCTCTCTCTCCACAGTCTC 2148  
Qy  
3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGCTCCCTCTCTCTCTCCACAGTCTC 3710  
Db  
2149 TAGGGCTGCGTAGCTGCAGGCTTCCAAAGGGGGTTTCAGTCTGGACTTATACAGGAGGC 2208  
Qy  
3711 TAGGGCTGCGTAGCTGCAGGCTTCCAAAGGGGGTTTCAGTCTGGACTTATACAGGAGGC 3770  
Db  
2209 CAGAAGGGCTCCATGCAGCTGGAATGCGGGGACTCTGCAGCTGGATTACCCAGGCTCAGGG 2268  
Qy  
3771 CAGAAGGGCTCCATGCAGCTGGAATGCGGGGACTCTGCAGCTGGATTACCCAGGCTCAGGG 3830  
Db  
2269 TTAACAGCTAGCTCCCTAGTTGAGACACACTAGAGAAAGGTTTGGGAGCTGAATAA 2328  
Qy  
3831 TTAACAGCTAGCTCCCTAGTTGAGACACACTAGAGAAAGGTTTGGGAGCTGAATAA 3890  
Db  
2329 CTCAGTCACCTGGTTTCCCATCTCTAGCCCTTAACCTGCAGCTTCGTTTAAATGAGCT 2388  
Qy  
3891 CTCAGTCACCTGGTTTCCCATCTCTAGCCCTTAACCTGCAGCTTCGTTTAAATGAGCT 3950  
Db  
2389 CTTGCATGGGAGTTTCTAGGATGAACACACTCCCTCCATGGGATTTGAACATATG--ACTTA 2446  
Qy  
3951 CTTGCATGGGAGTTTCTAGGATGAACACACTCCCTCCATGGGATTTGAACATATGAAAGTTA 4010  
Db  
2447 TTTGTAGGGAAGAGTCTCTAGGGGCAACACACAGAACAGTCCCTCAGCCACAGC 2506  
Qy  
4011 TTTGTAGGGAAGAGTCTCTAGGGGCAACACACAGAACAGTCCCTCAGCCACAGC 4070  
Db  
2507 ACTGTCTTTTGTGATCCACCCCTCTACCTTTATCAGGATGGCTGTGGTGC 2566  
Qy  
4071 ACTGTCTTTTGTGATCCACCCCTCTACCTTTATCAGGATGGCTGTGGTGC 4130  
Db  
2567 TTTCTGTGTCATCAGAGACAGGCAATTAATATTAATTTAACTATTTTAAACAAAGT 2626  
Qy  
4131 TTTCTGTGTCATCAGAGACAGGCAATTAATATTAATTTAACTATTTTAAACAAAGT 4190  
Db  
2627 AGAAGGAATCCATTCGTAGCTTTCTGTGTTGGTGTCTAAATTTGGTGGGTTGGGG 2686  
Qy  
4191 AGAAGGAATCCATTCGTAGCTTTCTGTGTTGGTGTCTAAATTTGGTGGGTTGGGG 4250  
Db  
2687 ATCCCAACAAATCAGTTCCTCAGATAGTGTGATGGCTGATGATCCAGATCT 2746  
Qy  
4251 ATCCCAACAAATCAGTTCCTCAGATAGTGTGATGGCTGATGATCCAGATCT 4310  
Db  
2747 TCTTCTCTGGGTCTGGCCCCCAAAATGCCCTAACCCAGGACCTTGGAAATTTACTCA 2806  
Qy  
4311 TCTTCTCTGGGTCTGGCCCCCAAAATGCCCTAACCCAGGACCTTGGAAATTTACTCA 4370  
Db  
2807 TCCCAATGATATTCCTCAATGCTGTACCAAGTTAGGTTAGGTTGAAGGAAGGTAGAG 2866  
Qy  
4371 TCCCAATGATATTCCTCAATGCTGTACCAAGTTAGGTTAGGTTGAAGGAAGGTAGAG 4430  
Db  
2867 GTGGGCTTCAGTCTCAAGGCTTCCCTAACCCCTCTCTAGGACTGGGTGATGAAGGACCTGCT 2926  
Qy  
4431 GTGGGCTTCAGTCTCAAGGCTTCCCTAACCCCTCTCTAGGACTGGGTGATGAAGGACCTGCT 4490  
Db  
2927 TCCCCCTTCCACTCCCTCTACTCTCTCTAGGACTGGGTGATGAAGGACCTGCTGCTT 2986  
Qy  
4491 TCCCCCTTCCACTCCCTCTACTCTCTCTAGGACTGGGTGATGAAGGACCTGCTGCTT 4550  
Db  
2987 AATTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 3046  
Qy  
4551 AATTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 4610  
Db  
3047 TGTGAGCTACTGAGGACAGAGCAAGTGGGTTTCCCAAGCTTTGTCATC 3106  
Qy  
4611 TGTGAGCTACTGAGGACAGAGCAAGTGGGTTTCCCAAGCTTTGTCATC 4670  
Db  
3107 TCAGCCCCCAGAGTATATCTGTGTTGGGAATCTCACAGAAACTCAGGAGCACCCTCC 3166  
Qy  
4671 TCAGCCCCCAGAGTATATCTGTGTTGGGAATCTCACAGAAACTCAGGAGCACCCTCC 4730  
Db

3167 TGCTCAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAAGTCCGCTTTCGAATAAG 3226  
Qy  
4731 TGCTCAGCTAAGGAGGCTTATCTCTCAGGGGGGTTTAAAGTCCGCTTTCGAATAAG 4790  
Db  
3227 TGCTCATTATTTATTTAGCGGGTGAATATTTATCTAGTGAAGCAATCAGAGTATAA 3286  
Qy  
4791 TGCTCATTATTTATTTAGCGGGTGAATATTTATCTAGTGAAGCAATCAGAGTATAA 4850  
Db  
3287 TGTTTATGTCGACAAATTTAAAGGCTTCTTATATGTTTTAAAAA 3330  
Qy  
4851 TGTTTATGTCGACAAATTTAAAGGCTTCTTATATGTTTTAAAAA 4894  
Db

## RESULT 15

AX267728 4894 bp DNA linear PAT 26-OCT-2001  
LOCUS  
DEFINITION Sequence 702 from Patent WO0173032.  
ACCESSION AX267728  
VERSION AX267728.1 GI:16516400  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,  
Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.  
and Henderson,R.A.  
TITLE Compositions and methods for the therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0173032-A 702 04-OCT-2001;  
CORTXA CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1..4894  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 928 a 1448 c 1354 g 1163 t 1 others  
ORIGIN

Query Match 62.8%; Score 2142.8; DB 6; Length 4894;  
Best Local Similarity 80.1%; Pred. No. 0;  
Matches 2872; Conservative 1; Mismatches 3; Indels 708; Gaps 2;  
Qy 455 GSCATTGTCCAGTCTGGGCTGTGCTGTGCTCCGCTCCTAGGCTCAGCCAGTGACCAC 514  
Db 1311 GSCATTGTCCAGTCTGGGCTGTGCTGTGCTCCGCTCCTAGGCTCAGCCAGTGACCAC 1370  
Qy 515 TGGCTGGACGCTATGGCCCGCCGCTTCTATCTGGGCACTGTCTTGGGCACTCCTG 574  
Db 1371 TGGCTGGACGCTATGGCCCGCCGCTTCTATCTGGGCACTGTCTTGGGCACTCCTG 1430  
Qy 575 CTGAGCCCTTCTCATCCCAAGGCGCGCTGTGTAGAGGGCTGTGTGCCCCGATCCC 634  
Db 1431 CTGAGCCCTTCTCATCCCAAGGCGCGCTGTGTAGAGGGCTGTGTGCCCCGATCCC 1490  
Qy 635 AGGCCCTTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGTGTGCCCCGATCCC 694  
Db 1491 AGGCCCTTGGAGCTGGCACTGCTCATCTGGGCGTGGGCTGTGTGCCCCGATCCC 1550  
Qy 695 GTGTGCTTCTACTCCACTGGAGGCTTGTCTTCTGACCTCTTCCCGGACCCGAGCACTGT 754  
Db 1551 GTGTGCTTCTACTCCACTGGAGGCTTGTCTTCTGACCTCTTCCCGGACCCGAGCACTGT 1610  
Qy 755 CGCCAGGCTTACTGTCTATGCGCTTATGATAGTCTTGGGGCTGTGTGGGCTACCTC 814  
Db 1611 CGCCAGGCTTACTGTCTATGCGCTTATGATAGTCTTGGGGCTGTGTGGGCTACCTC 1670  
Qy 815 CTGCTGCAATGTAGTGGACACAGTGCCCTGCCCCCTACCTGGGCACTTGGGCACTGAG 874  
Db 1671 CTGCTGCAATGTAGTGGACACAGTGCCCTGCCCCCTACCTGGGCACTTGGGCACTGAG 1730

Qy	575	CTGAGGCTCTTTCTCATGCCAAGGCGCGGCTGGCTAGCAGGCGTCTGTCGCCCGATCCC	634
Db	1431	CTGAGGCTCTTTCTCATGCCAAGGCGCGGCTGGCTAGCAGGCGTCTGTCGCCCGATCCC	1490
Qy	635	AGGCCCTCGAGCTGGCACTGCTCATCTCGGGCGTGGGCTGCTGGACTTCTGTGGCCAG	694
Db	1491	AGGCCCTCGAGCTGGCACTGCTCATCTCGGGCGTGGGCTGCTGGACTTCTGTGGCCAG	1550
Qy	695	GTGTGCTTCACATCCACTGAGGCGCTGCTCTGACCTTTCGGGGACCCGGACCACTGT	754
Db	1551	GTGTGCTTCACATCCACTGAGGCGCTGCTCTGACCTTTCGGGGACCCGGACCACTGT	1610
Qy	755	CGCCAGCGCTACTCTGCTATGCCTTCATGATCAGTCTTGGGGGCTGCTGGGCTACCTC	814
Db	1611	CGCCAGCGCTACTCTGCTATGCCTTCATGATCAGTCTTGGGGGCTGCTGGGCTACCTC	1670
Qy	815	CTGCCTGCCATTTGACTGGGACACCACTGCGCTGGCCCCCTACTGGGCACCCAGAGGAG	874
Db	1671	CTGCCTGCCATTTGACTGGGACACCACTGCGCTGGCCCCCTACTGGGCACCCAGAGGAG	1733
Qy	875	TGCGCTTTTGGCGCTGCTACCGCTCATCTTCCTCACTGCGTAGCAGGCACACTGCTGGT	934
Db	1731	TGCGCTTTTGGCGCTGCTACCGCTCATCTTCCTCACTGCGTAGCAGGCACACTGCTGGT	1790
Qy	935	GCTGAGGAGCAGCGCTGGGCGCCCCACCGACGACGAGGCGCTGTCGGCCCCCTCTTG	994
Db	1791	GCTGAGGAGCAGCGCTGGGCGCCCCACCGACGACGAGGCGCTGTCGGCCCCCTCTTG	1850
Qy	995	TCGCGCCACTGTGTCATGCGCGGGCCGGTTGGCTTTCGGAACCTGGCGGCCCTGCTT	1054
Db	1851	TCGCGCCACTGTGTCATGCGCGGGCCGGTTGGCTTTCGGAACCTGGCGGCCCTGCTT	1910
Qy	1055	CCCGGGCTGACACAGCTGTGCTGCCGATGCCCGCACCGCTGCGCGGCTCTTCGTGGCT	1114
Db	1911	CCCGGGCTGACACAGCTGTGCTGCCGATGCCCGCACCGCTGCGCGGCTCTTCGTGGCT	1970
Qy	1115	GAGCTGTGACGTGGATGGCACTCATGACCTTCACGCTGTTTTACAGGATTTCTGTGGC	1174
Db	1971	GAGCTGTGACGTGGATGGCACTCATGACCTTCACGCTGTTTTACAGGATTTCTGTGGC	2030
Qy	1175	GAGGGGCTGTACAGGGCGTGCCACAGCTGAGCGCGGCACCGAGCGCGGACACTAT	1234
Db	2031	GAGGGGCTGTACAGGGCGTGCCACAGCTGAGCGCGGCACCGAGCGCGGACACTAT	2090
Qy	1235	GATGA-----	1239
Db	2091	GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGTGGAGCGCCGCCACAGAGAGC	2150
Qy	1240	-----	1239
Db	2151	ACACTCGGGCTGTGTCTGGGCTGGTGCCCTCTCCATCCTGGCCCGACTTCTGTGACG	2210
Qy	1240	-----	1239
Db	2211	AAAGTGGGATGGACCCCATCTGCATACAGGCTTCTCATGGGTGTGGAACATCTCTGCT	2270
Qy	1240	-----	1239
Db	2271	TSCGGTTTCAGGAAGGCTCTGGGCTGTCTAGGAGTCTGATCAGAGTCGTTGCCCGATT	2330
Qy	1240	-----	1239
Db	2331	TGACAGAAGGAAGGGGAGCTTATTCAAAGCTCTAGAGGAGTGGAGGAGTTAAGGCTGG	2390
Qy	1240	-----	1239
Db	2391	ATTTTCAGATCTGCCTGGTTCCAGCGCAGTGTGCCCTCTGCTCCCCAAGGACTTTCCAA	2450
Qy	1240	-----	1239
Db	2451	ATAATCTCACAGCGCGCTTCCAGCTCAGGCGTCTTAGAAGGCTTTGAAGCTATGGCCA	2510
Qy	1240	-----	1239

Db	2511	GC	TGCTTTGTGTTGTCCTCTACCCGCCCTGTCTCTACAGCTCAGACTCCAGGAAACCTT	2570	
QY	1240	-	-	1239	
Db	2571	CAGACTACCTT	CCTCTGCTT	CAGCAAGGGCGTTGCCACATTTCTCTGAGGTCAGTGG	2630
QY	1240	-	-	1239	
Db	2631	AAGAACTAGACTTCCCATTTGCTAGAGGTAGAAAGGGAGGGTGTCTGGGAGCAGGGCTG	2690		
QY	1240	-	-	1239	
Db	2691	GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCCTTCTCATCTCCTCGAGAC	2750		
QY	1240	-	-	1248	
Db	2751	TGCTCCGACCTTTCCTCCAGGCTGTGTATGAGCCCTCTCCCTCTGCAGGCTTTGG	2810		
QY	1249	GATGGGAGCCTGGGGCTGTTTCTTCAGTCAGTCGCCCATCTCCCTGGCTCTCTCTCTGTCAT	1308		
Db	2811	GATGGGAGCCTGGGGCTGTCTCTGAGTCGCCATCTCCCTGGTCTTCTCTCTGTCAT	2870		
QY	1309	GGACCGGCTGGTGCAGCGATTTCGGCACTCGAGCAGTCTATTGGCCAGTGTGGCAGCTTT	1368		
Db	2871	GGACCGGCTGGTGCAGCGATTTCGGCACTCGAGCAGTCTATTGGCCAGTGTGGCAGCTTT	2930		
QY	1369	CCCTGTGGCTGCCGGTGCCACATGCTGTGCCACAGTGTGGCCGTGGTGACAGCTTCAGC	1428		
Db	2931	CCCTGTGGCTGCCGGTGCCACATGCTGTGCCACAGTGTGGCCGTGGTGACAGCTTCAGC	2990		
QY	1429	CGCCCTCACCGGTTACCTTCTCAGCCCTGCAGATCTCCGCTACACACTGGCCCTCCCT	1488		
Db	2991	CGCCCTCACCGGTTACCTTCTCAGCCCTGCAGATCTCCGCTACACACTGGCCCTCCCT	3050		
QY	1489	CTACCAACCGGGAAGCAGGTGTTCTTGCCCAATACCGAGGGACACTGGAGGTGCTAG	1548		
Db	3051	CTACCAACCGGGAAGCAGGTGTTCTTGCCCAATACCGAGGGACACTGGAGGTGCTAG	3110		
QY	1549	CAGTGAGCAGCCTGATGACCAGCTTCTTGCCAGGCCCTTAAGCCTGGAGCTCCCTTCCC	1608		
Db	3111	CAGTGAGCAGCCTGATGACCAGCTTCTTGCCAGGCCCTTAAGCCTGGAGCTCCCTTCCC	3170		
QY	1609	TAATGGACACTGGGTGCTGGAGCAGTGGCTGCTCCCACTCCACCCGCGTCTCGG	1668		
Db	3171	TAATGGACACTGGGTGCTGGAGCAGTGGCTGCTCCCACTCCACCCGCGTCTCGG	3230		
QY	1669	GGCCTCTGCCTGTGATGTCCTCGTACGTGTGGTGGTGAGCCCAACGAGGCCAGGGT	1728		
Db	3231	GGCCTCTGCCTGTGATGTCCTCGTACGTGTGGTGGTGAGCCCAACGAGGCCAGGGT	3290		
QY	1729	GGTTCCGGGGCGGGCATCTGCCCTGGACCTCGCCATCTGGATAGTGCCTTCTGCTGTC	1788		
Db	3291	GGTTCCGGGGCGGGCATCTGCCCTGGACCTCGCCATCTGGATAGTGCCTTCTGCTGTC	3350		
QY	1789	CCAGGTGGCCCATCTCCCTGTTTATGGGCTCCATTTGCTCAGCTCAGCCAGTCTGTCACTGC	1848		
Db	3351	CCAGGTGGCCCATCTCCCTGTTTATGGGCTCCATTTGCTCAGCTCAGCCAGTCTGTCACTGC	3410		
QY	1849	CTATATGTTGTCCTGCCAGCCCTGGGTCTGGTTCGCCATTTTACCTTGTGTACACAGGTAGT	1908		
Db	3411	CTATATGTTGTCCTGCCAGCCCTGGGTCTGGTTCGCCATTTTACCTTGTGTACACAGGTAGT	3470		
QY	1909	ATTTGACAAGACGACTTGGCCAAATACTCAGCTAGAAACTTCCAGCACATTTGGGTG	1968		
Db	3471	ATTTGACAAGACGACTTGGCCAAATACTCAGCTAGAAACTTCCAGCACATTTGGGTG	3530		
QY	1969	GAGGCGCTGCCTCACTGGGTCCCAGCTCCCGCTCCTGTTAGCCCCCATGGGCTGCCGG	2028		
Db	3531	GAGGCGCTGCCTCACTGGGTCCCAGCTCCCGCTCCTGTTAGCCCCCATGGGCTGCCGG	3590		
QY	2029	CTGGCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCTGTGCTG	2088		

Db	1321	GCCTGGACCTCGCCATCCTGGATAGTGGCTTCCTGCTGCCAGGTGGCCCATCCCTGT	1380
QY	1809	TTATGGGCTCAATTTGTCAGCTCAGCAGTCTGTCACTGCCTATATGTTGTTGCTGCGCAG	1868
Db	1381	TTATGGGCTCCATTTGTCAGCTCAGCAGTCTGTCACTGCCTATATGTTGTTGCTGCGCAG	1440
QY	1869	GCCTGGGTCGTGGTCCCATTTACTTTGCTACAGAGTAGTATTTGACAAAGACGACTGG	1928
Db	1441	GCCTGGGTCGTGGTCCCATTTACTTTGCTACAGAGTAGTATTTGACAAAGACGACTGG	1500
QY	1929	CCAATATCTCAGCGTAGAAAACCTCCAGCACATTTGGGFTGGAGGCCCTGCCTCACTGGGT	1988
Db	1501	CCAATATCTCAGCGTAGAAAACCTCCAGCACATTTGGGFTGGAGGCCCTGCCTCACTGGGT	1560
QY	1989	CCAGCTCCCGGCTCCTGTTAGCCCCATGGGGTGGCGGCTGGCGCCAGTTCTTGTGTTG	2048
Db	1561	CCAGCTCCCGGCTCCTGTTAGCCCCATGGGGTGGCGGCTGGCGCCAGTTCTTGTGTTG	1620
QY	2049	CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCTGTGCTGTGAGGTGCTGTAGCTGCACA	2108
Db	1621	CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCTGTGCTGTGAGGTGCTGTAGCTGCACA	1680
QY	2109	GCTGGGGCTGGGGCTGCCCTCTCTCTCTCCTCCAGTCTCTAGGCTGCCCTGACTGGAGG	2168
Db	1681	GCTGGGGCTGGGGCTGCCCTCTCTCTCTCCTCCAGTCTCTAGGCTGCCCTGACTGGAGG	1740
QY	2169	CCTTCCAAAGGGGTTTCAGTCTGAGCTTATACAGAGGCCAGAGGGCTCCATGCACCTG	2228
Db	1741	CCTTCCAAAGGGGTTTCAGTCTGAGCTTATACAGAGGCCAGAGGGCTCCATGCACCTG	1800
QY	2229	GAATGCGGGGACTCTGCAGGTGGATTACCAGGCTCAGGGTTAACAGTAGACCTCCCTAGT	2288
Db	1801	GAATGCGGGGACTCTGCAGGTGGATTACCAGGCTCAGGGTTAACAGTAGACCTCCCTAGT	1860
QY	2289	TGAGACACCTAGAGAAGGGTTTTGGGAGCTGAATAAATCAGTCAGCTCAGCTGGTTCCCA	2348
Db	1861	TGAGACACCTAGAGAAGGGTTTTGGGAGCTGAATAAATCAGTCAGCTCAGCTGGTTCCCA	1920
QY	2349	TCTCTAAGCCCTTAACTGCAGCTTCGTTTAAATAGCTCTCTGCATGGGAGTTTCFAGG	2408
Db	1921	TCTCTAAGCCCTTAACTGCAGCTTCGTTTAAATAGCTCTCTGCATGGGAGTTTCFAGG	1980
QY	2409	ATGAACACTCTCCATGGGATTTGAACATATG--ACTTATTTAGGGGAAGAGTCCCTG	2466
Db	1981	ATGAACACTCTCCATGGGATTTGAACATATGAAAGTTATTTAGGGGAAGAGTCCCTG	2040
QY	2467	AGGGGAACACACAAGAACCCAGTCCCTCAGCCCCACAGCACTGCTTTTTGCTGATCCA	2526
Db	2041	AGGGGAACACACAAGAACCCAGTCCCTCAGCCCCACAGCACTGCTTTTTGCTGATCCA	2100
QY	2527	CCCCCTCTTACTTTATCAGAGTGGGCTGTGTCCTTCTGTTGCCATCACAGAGA	2586
Db	2101	CCCCCTCTTACTTTATCAGAGTGGGCTGTGTCCTTCTGTTGCCATCACAGAGA	2160
QY	2587	CACAGCATTTAAATATTTAACTATTTATTAACAAGTAGAAGGAATCCATTCGTAG	2646
Db	2161	CACAGCATTTAAATATTTAACTATTTATTAACAAGTAGAAGGAATCCATTCGTAG	2220
QY	2647	CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGTTGGGGATCCCAACATCAGGTCCC	2706
Db	2221	CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGTTGGGGATCCCAACATCAGGTCCC	2280
QY	2707	CTGAGATAGCTGGTCATTTGGGCTGATCATTCGAGAAATCTTCTCTCTGGGCTGGGC	2766
Db	2281	CTGAGATAGCTGGTCATTTGGGCTGATCATTCGAGAAATCTTCTCTCTGGGCTGGGC	2340
QY	2767	CCCCAAATAGCCTAAACCCAGAGACCTTGGAAATTTCTACTCATCCCAATGATATCCAAA	2826
Db	2341	CCCCAAATAGCCTAAACCCAGAGACCTTGGAAATTTCTACTCATCCCAATGATATCCAAA	2400
QY	2827	TGCTGTATCCCAAGGTTTAGGTTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGCTCTCAAC	2886
Db	2401	TGCTGTATCCCAAGGTTTAGGTTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGCTCTCAAC	2466
QY	2887	GGCTTCCCTTAACCAACCCCTCTTCTCTTGGGCCAGCCTGTTTCCGCCCACTTCACATCCCC	2946
Db	2461	GGCTTCCCTTAACCAACCCCTCTTCTCTTGGGCCAGCCTGTTTCCGCCCACTTCACATCCCC	2520
QY	2947	TCTACTCTCTCTAGGACTGGGCTGATGAAGCACTGCCCAAAATTTTCCCTACCCCAAC	3006
Db	2521	TCTACTCTCTCTAGGACTGGGCTGATGAAGCACTGCCCAAAATTTTCCCTACCCCAAC	2580
QY	3007	TTTCCGCTACCCCAACCTTTCCCAACAGCTCCCAACACCTGTTTGGAGCTACTGCAGGA	3066
Db	2581	TTTCCGCTACCCCAACCTTTCCCAACAGCTCCCAACACCTGTTTGGAGCTACTGCAGGA	2640
QY	3067	CCAGAAGCACAAAGTGCAGTTTCCCAAGCCTTTGCTCCATCTCAGCCCCCAGAGTATATCT	3126
Db	2641	CCAGAAGCACAAAGTGCAGTTTCCCAAGCCTTTGCTCCATCTCAGCCCCCAGAGTATATCT	2700
QY	3127	GTGCTTGGGGAATCTCACAGAACTCAGAGCACCCCTGCCTGAGCTAAGGGAGGTC	3186
Db	2701	GTGCTTGGGGAATCTCACAGAACTCAGAGCACCCCTGCCTGAGCTAAGGGAGGTC	2760
QY	3187	TTATCTCTCAGGGGGGTTTAAAGTCCGTTTGAATAATGCTGCTTATTTATTTAGCGG	3246
Db	2761	TTATCTCTCAGGGGGGTTTAAAGTCCGTTTGAATAATGCTGCTTATTTATTTAGCGG	2820
QY	3247	GCTGAATATTTTATCTAGTGAAGTGAAGCAATCAGAGTATAATGTTTGTGCAAAATTA	3306
Db	2821	GCTGAATATTTTATCTAGTGAAGTGAAGCAATCAGAGTATAATGTTTGTGCAAAATTA	2880
QY	3307	RAGGCTTTCTTATATGTTTAAAAA 3330	
Db	2881	RAGGCTTTCTTATATGTTTAAAAA 2904	
RESULT 14			
AX200993			
LOCUS	AX200993	4894 bp	DNA linear PAT 29-AUG-2001
DEFINITION	Sequence 623 from Patent WO0151633.		
ACCESSION	AX200993		
VERSION	AX200993.1	GI:15390820	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 4894)		
AUTHORS	Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,X., Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W., Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.		
TITLE	Compositions and methods for the therapy and diagnosis of prostate cancer		
JOURNAL	Patent: WO 0151633-A 623 19-JUL-2001;		
FEATURES	CORIXA CORPORATION (US)		
source	Location/Qualifiers		
	1..4894		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	928 a 1448 c 1354 g 1163 t	1 others	



QY 3007 TTTCCTTACCCCACTTTCCACCAGCTCCACACCCCTGTTGGAGCTACTGCAGGA 3066  
Db 2581 TTTCCTTACCCCACTTTCCACCAGCTCCACACCCCTGTTGGAGCTACTGCAGGA 2640  
QY 3067 CCAGAGCAAAAGTGCCTTTCCCAAGCTTTGTCATCTCAGCCGCCAGAGTATATCT 3126  
Db 2641 CCAGAGCAAAAGTGCCTTTCCCAAGCTTTGTCATCTCAGCCGCCAGAGTATATCT 2700  
QY 3127 GTGCTTGGGAATCTCACACAGAACTCAGAGACACCCCTGCTGAGCTAAGGAGGTC 3186  
Db 2701 GTGCTTGGGAATCTCACACAGAACTCAGAGACACCCCTGCTGAGCTAAGGAGGTC 2760  
QY 3187 TTATCTCTCAGGGGGGTTAAGTGCCTTTGCAATATGCTGCTTATTATTATAGCGG 3246  
Db 2761 TTATCTCTCAGGGGGGTTAAGTGCCTTTGCAATATGCTGCTTATTATTATAGCGG 2820  
QY 3247 GGTGAATATTTATCTGTAAGTGCAGCAATCAGAGTATATGTTATGTTGACAAAATTA 3306  
Db 2821 GGTGAATATTTATCTGTAAGTGCAGCAATCAGAGTATATGTTATGTTGACAAAATTA 2880  
QY 3307 AAGGCTTCTTATATGTTTAAAAA 3330  
Db 2881 AAGGCTTCTTATATGTTTAAAAA 2904

RESULT 13  
AX267729  
LOCUS AX267729 2904 bp DNA linear PAT 26-OCT-2001  
DEFINITION Sequence 703 from Patent WO0173032.  
ACCESSION AX267729  
VERSION AX267729.1 GI:16516401  
KEYWORDS  
SOURCE human  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
XU, J., DILLON, D. C., MITCHAM, J. L., HARLOCKER, S. L., JIANG, Y.,  
KALOS, M. D., FANGER, G. R., RETTER, M. W., STOLK, J. A., DAY, C. H.,  
VEDVICK, T. S., CARTER, D., LI, S. X., WANG, A., SKEIKY, Y. A., HEPLER, W. T.  
and HENDERSON, R. A.  
Compositions and methods for the therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0173032-A 703 04-OCT-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1..2904  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 542 a 875 c 773 g 714 t  
ORIGIN

Query Match 64.4%; Score 2196.4; DB 6; Length 2904;  
Best Local Similarity 88.2%; Pred. No. 0;  
Matches 2560; Conservative 0; Mismatches 1; Indels 343; Gaps 2;

QY 770 GTCTATGCTTCATGATCAGTCTTTGGGGGCTGCTGGGCTACCTCTGCTGCCATTGAC 829  
Db 1 GTCTATGCTTCATGATCAGTCTTTGGGGGCTGCTGGGCTACCTCTGCTGCCATTGAC 60  
QY 830 TGGGACACAGTGGCTGGCCCCCTACCTGGGACACCCAGGAGGAGTGCCTCTTTGGCCTG 889  
Db 61 TGGGACACAGTGGCTGGCCCCCTACCTGGGACACCCAGGAGGAGTGCCTCTTTGGCCTG 120  
QY 890 CTCACCTTCATCTTCCTACCTGCTAGCAGCCACACTGCTGGGCTGAGGAGGAGCGG 949  
Db 121 CTCACCTTCATCTTCCTACCTGCTAGCAGCCACACTGCTGGGCTGAGGAGGAGCGG 180  
QY 950 CTGGGCCCCAGCGACCGAGAGAGGGTGTGCGCCCCCTCTTGTGCCCCCACTGCTGT 1009  
Db 181 CTGGGCCCCAGCGACCGAGAGAGGGTGTGCGCCCCCTCTTGTGCCCCCACTGCTGT 240

QY 1010 CCATGCGCGCGCCCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCGGGCTGCACAG 1069  
Db 241 CCATGCGCGCGCCCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCGGGCTGCACAG 300  
QY 1070 CTGTGCTGCGCATGCCCCCGCACCCCTGCGCGGCTCTTCTGCTGAGCTGTGACGTGG 1129  
Db 301 CTGTGCTGCGCATGCCCCCGCACCCCTGCGCGGCTCTTCTGCTGAGCTGTGACGTGG 360  
QY 1130 ATGGCACTCATGACCTTTCACGCTCTTTTACACGATTTCCTGGGCGAGGGGCTGTACCA 1189  
Db 361 ATGGCACTCATGACCTTTCACGCTCTTTTACACGATTTCCTGGGCGAGGGGCTGTACCA 420  
QY 1190 GCGTGGCCAGAGCTGAGCGGGCCAGAGCCCGGAGACACATATATGATGAA 1240  
Db 421 GCGTGGCCAGAGCTGAGCGGGCCAGAGCCCGGAGACACATATGATGAAAGAGCCT 480  
QY 1241 1241 1240  
Db 481 CTGCTGCTCTAGGAGTCTGATCAGAGTCTGTTGCCCAAGTTTGACAGAAGAAAGCGGA 540  
QY 1241 1241 1240  
Db 541 GCTTATTCAAAGTCTAGAGGAGTGGAGAGTAAAGGCTGGATTTCAGATCTGCTGTT 600  
QY 1241 1241 1240  
Db 601 CCAGCGCGAGTGTGCTCTGCTCCCAACAGACTTTTCCAAATATCTCACCAGCGCTT 660  
QY 1241 1241 1240  
Db 661 CCAGCTCAGGCGTCTAGAGCGTCTTGAAGCCTATGCGCAGCTGTCTTGTGTTCCCTC 720  
QY 1241 1241 1240  
Db 721 TCACCCGCTGTGCTCAGAGCTGAGACTCCAGGAAACCTTTCAGACTACCTTCTCTGCC 780  
QY 1241 1241 1268  
Db 781 TTCACAGAGGGCGTGTGCCACATCTCTGAGGGGCTTCGATGGGAGCCTCGGGCTGT 840  
QY 1269 TCCTCAGTGGCGCATCTCCCTGTGCTCTCTCTGCTCATGACCGGCTGTGACGAT 1328  
Db 841 TCCTCAGTGGCGCATCTCCCTGTGCTCTCTCTGCTCATGACCGGCTGTGACGAT 900  
QY 1329 TCGGCATCGAGCAGTCTATTGCGCAGTGTGGCAGCTTCCCTGTGCTGCCGTGCCA 1388  
Db 901 TCGGCATCGAGCAGTCTATTGCGCAGTGTGGCAGCTTCCCTGTGCTGCCGTGCCA 960  
QY 1389 CATGCTGTCCACAGTGTGGCGCTGTGACAGCTTCAGCGCCCTCACCAGGTTACCT 1448  
Db 961 CATGCTGTCCACAGTGTGGCGCTGTGACAGCTTCAGCGCCCTCACCAGGTTACCT 1020  
QY 1449 TCTAGCCCTGCAGATCTGCTTACACATGCTGCTTCTTACACCGGAGAGAGG 1508  
Db 1021 TCTAGCCCTGCAGATCTGCTTACACATGCTGCTTCTTACACCGGAGAGAGG 1080  
QY 1509 TGTTCCTGCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGCAGACAGCTGATGA 1568  
Db 1081 TGTTCCTGCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGCAGACAGCTGATGA 1140  
QY 1569 CCAGCTTCTGCGCAGCCCTTAAGCCTTGGAGCTTCCCTTCCCTAATGGACAGCTGGTGTG 1628  
Db 1141 CCAGCTTCTGCGCAGCCCTTAAGCCTTGGAGCTTCCCTTCCCTAATGGACAGCTGGTGTG 1200  
QY 1629 GAGGAGTGGCTGTCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCTGTGATGCT 1688  
Db 1201 GAGGAGTGGCTGTCTCCACCTCCACCGCGCTCTGCGGGGCTCTGCTGTGATGCT 1260  
QY 1689 CCGTACGTGTGTGGTGGGTGAGCCACCGAGGAGGTGCTCCGGCCGCGCATCT 1748  
Db 1261 CCGTACGTGTGTGGTGGGTGAGCCACCGAGGAGGTGCTCCGGCCGCGCATCT 1320  
QY 1749 GCCTGGACCTCGCCATCTCTGGATGATGCTTCTGCTGCCAGGTTGCCCATCCCTGT 1808



